Gyires's permanent inequality for structured matrices

Dev D. Goyal

October 2026; revised 1 November 2026

Abstract

Gyires's conjecture [10], listed as Conjecture 20 by Cheon and Wanless [5], compares the square of the permanent of a doubly stochastic matrix with three permanents of matrix products. This paper proves the inequality for several structured classes of matrices. These classes include nonnegative circulants of orders five, six, and seven; all-order permutation polynomials $aI+bQ+cQ^2$ with nonnegative coefficients; all-order doubly stochastic matrices with at most two row types and two column types; matrices with one exceptional column or row; and all rank-at-most-two doubly stochastic matrices through order fifteen. The circulant results include sharp sorted-gap margins and equality cases. Further results treat sparse perturbations, direct sums, positive-semidefinite matrices, and matrix boundary reduction. The supporting arguments include scalar norm inequalities, polynomial deletion, balanced limits, local stable and Lorentzian inequalities, and exact configuration-model identities. Exact counterexamples identify the limits of proposed extensions. In Lean 4, the principal order-five and order-six claims, the all-order permutation-polynomial, four-block and exceptional-vector theorems, and the rank-two theorem for orders 2 through 8 have complete proofs, and the general permanent comparisons and direct-sum closure also have Lean proofs; the order-seven circulant theorem and rank-two orders 9 through 15 rest on conventional proofs with exact certificates and remain incomplete in Lean. The unrestricted conjecture is not resolved.

Contents

1 The problem and the results **6**

1.1 Gyires's inequality 7

1.2 Matrix results and proof methods 7

1.3 Supporting results and their limits 7

1.4 Verification status 8

1.5 Prior work 8

2 Common permanent identities **8**

3 Nonnegative circulants of order five **9**

3.1 Notation 10

3.2 The inequalities and sharp margin 10

3.3 Certificate method 11

3.4 Affine symmetry 11

3.5 Six chambers . . . 12
3.6 The chamber certificate . . . 12
3.7 Proof of Theorems 3.5 and 3.4 . . . 13
3.8 Proof of Corollary 3.6 . . . 13

4 Nonnegative circulants of orders six and seven 13
4.1 Affine symmetry and complete coverage . . . 14
4.2 Construction and independent checking of the polynomials . . . 15
4.3 Sharpness, equality, and the Gyires deduction . . . 16
4.4 A consequence in unbounded dimension . . . 17
4.5 Reproduction and limits . . . 17

5 Permutation polynomials in arbitrary dimensions 18
5.1 The cyclic permanent and its transfer matrix . . . 18
5.2 A uniform spectral bound . . . 20
5.3 A weighted trace bound . . . 21
5.4 The analytic bound from order 23 . . . 24
5.5 An additional analytic bound from order 48 . . . 24
5.6 Exact polynomial certificates for orders 3 through 22 . . . 25
5.7 Small orders, boundary cases, and permutation cycles . . . 27

6 Four-block matrices in arbitrary dimensions 28
6.1 The four blocks and their full parameter interval . . . 28
6.2 A probability inequality at an integer mean . . . 29
6.3 From two column types to a permanent contraction . . . 32
6.4 Proof of the four-block theorem . . . 33

7 Matrices with one exceptional column or row 34
7.1 Statement . . . 34
7.2 Parametrization and the first Gram inequality . . . 34
7.3 An exact formula for the other Gram permanent . . . 34
7.4 Two elementary lower bounds for the one-success probability . . . 35
7.5 The range $1/n \le a \le 1/2$ . . . 36
7.6 The range $1/2 \le a \le 1$ . . . 36
7.7 Equality and Gyires's inequality . . . 37

8 All rank-two matrices in orders two through fifteen 37
8.1 A representation of every rank-two matrix . . . 37
8.2 Separate convexity in the probability laws . . . 38

8.3 Complete coverage of the profiles and endpoints . . . 40
8.4 Homogeneous integer certificates . . . 40
8.5 The exhaustive finite calculation . . . 41
8.6 Conclusion and equality . . . 42

9 Closure, sparse perturbations, and boundary reduction 43
9.1 Direct sums and independent permutations . . . 43
9.2 Positive-semidefinite matrices . . . 44
9.3 Three exceptional rows and an exact cubic expansion . . . 44
9.4 A trace bound for centered positive-semidefinite perturbations . . . 46
9.5 The symmetrized denominator along a radial segment . . . 48
9.6 A direct reduction to the matrix boundary . . . 49

10 Scalar and two-group polynomial inequalities 50
10.1 Localization of likelihood ratios . . . 50
10.2 The mean-one binomial theorem . . . 51
10.2.1 The balancing tilt . . . 52
10.2.2 Comparison with the omitted binomial tail . . . 53
10.2.3 Equality . . . 54
10.2.4 Polynomial consequence . . . 54
10.3 Symmetric reference laws . . . 55
10.3.1 Statement . . . 55
10.3.2 Proof . . . 55
10.3.3 Central binomial consequence . . . 56
10.3.4 Exact examples . . . 56
10.3.5 Two-group polynomial consequence . . . 57
10.4 Duplication and the Gram permanent . . . 57
10.4.1 Scalar bridge theorem . . . 58
10.4.2 Proof . . . 58
10.4.3 Polynomial form . . . 59
10.5 The remaining general integer-mean reduction . . . 59

11 A two-law scalar reduction in arbitrary order 60
11.1 The permanent and a binomial-noise operator . . . 60
11.2 The larger class and its variance bound . . . 61
11.3 Separate convexity of the square-root deficit . . . 61
11.4 Localization to two support intervals . . . 62

12 Coordinate deletion and exact finite certificates **63**

12.1 The cubic deletion theorem 64

12.2 Birkhoff parametrization and chamber coverage 64

12.3 Independent exact verification 65

12.4 Equality in the order-three certificate 66

12.5 Conditional induction: normalization and closure 66

12.6 Boundary cases can be handled without assuming finite balancing 67

12.7 The permanent-gradient identity 68

12.8 Deletion on a two-shift cycle 68

13 Rank-two deletion and balanced limits **70**

13.1 Extraction of one copy of a variable 70

13.2 Capacity and a unique balanced limit 71

13.3 The exact missing step for induction 73

13.4 A real-rooted collision marginal 74

14 Local polynomial inequalities and boundary reductions **75**

14.1 The polynomial classes and the deletion functional 75

14.2 The stable tangent cone and the sharp local theorem 77

14.3 A variance-deficit quantity and exact local control 80

14.4 The Lorentzian local theorem 81

14.5 Why balance alone does not suffice 82

14.6 Subtraction to a pure-coefficient boundary 83

14.7 Cubic stability, deletion, and boundary rigidity 84

14.8 An admissible attenuation and a finite support identity 86

14.9 A multivariate obstruction to two-sided scalar localization 87

15 Exact expectations in a regular bipartite configuration model **89**

15.1 Counting proof 89

15.2 The second moment 90

16 Exact limits of stronger proof steps **90**

16.1 Negative coefficients in a general Birkhoff sorting chamber 91

16.2 A zero-containing column need not be a useful deletion 91

16.3 The ultra-log-concave class does not satisfy the stronger contraction 93

16.4 Stochastic complements require more than a pivot sign 94

16.5 Averaging minimum norm scales does not replace capacity balancing 95

16.6 Admissible attenuation need not improve the balanced norm 97

17 Verification and reproduction **99**
17.1 Formal rank-two theorem in orders two through eight 99
17.2 Order-five coefficient checks 99
17.3 Formal certificate for the order-six theorem 100
17.4 Permutation-polynomial Lean correspondence 100
17.5 The four-block argument 101
17.6 Four-block Lean correspondence 102
17.7 Exceptional-vector Lean correspondence 102
17.8 Order-five Lean correspondence 102
17.9 General supporting Lean results 104
17.10 Limits of the formal verification 104

18 Comparison of the results and remaining questions **104**
18.1 A two-point example and a distinction between the bounds 105
18.2 The elementary order-two case 105
18.3 What remains outside the proofs 105

19 Research process and AI disclosure **105**

A Additional results for permanent and polynomial inequalities **106**
A.1 Definitions 106
A.2 The central coefficient need not dominate repeated columns 107
A.3 Pairwise averaging need not decrease the strong ratio 107
A.4 Normalized coefficients are not multivariate Jensen-concave 108
A.5 The collision polynomial can lose stability and balance 108
A.6 A global exponential trace lower bound fails 109
A.7 Fourier magnitudes and a fixed transfer gap are insufficient 110
A.8 Two stability failures for circulant coefficient constructions 110
A.9 Transpose averaging does not restore stability or collision balance 111
A.10 An operator inequality stronger than the scalar target fails 112
A.11 Positive-semidefinite completion does not give coefficientwise positivity 112
A.12 Matrix symmetrization and signed contractions do not give the required reductions 113
A.13 The occupancy entropy bridge fails with the actual probabilities 114
A.14 Balance and negative dependence do not replace product structure 115
A.15 A mixed-coefficient estimate fails for Lorentzian cubics 115
A.16 Three limits of scalar likelihood reductions 116
A.17 A collision probability need not decrease with the trial counts 117
A.18 Column averaging can increase a rank-two permanent 118

A.19 Minimum Fischer norm cannot supply a proposed deletion factor 118
A.20 An entropy reduction for scalar reference laws 119
A.21 An equivalent product inequality 120
A.22 Tensor forms of the two inequalities 121
A.23 The first two gap layers for Latin-square weight matrices 121
A.24 A two-product relaxation for balanced stable ternary cubics 122
A.25 Two explicit constructions within polynomial deletion 123
A.26 A Hessian criterion and a family of Lorentzian coefficients 124
A.27 The two Gram permanents need not have a fixed order 125
A.28 Two alternative bounds for the three-shift transfer matrices 125

B Finite computational observations **127**

B.1 Exact coefficient tests outside the complete circulant orders 128
B.2 Wider finite checks for the three-shift family 128
B.3 Matrix ratio searches 129
B.4 Circulant ratio and Fourier-phase searches 129
B.5 Configuration-model expectations 130
B.6 Stable-polynomial and matroid deletion tests 131
B.7 Circulant deletion tests 132
B.8 Nested singleton deletion and concentrated elementary-symmetric inputs 132
B.9 Direct Lorentzian coefficient optimization 133
B.10 M-convex support enumeration and attenuation 134
B.11 Rank-two scalar and deletion tests 134
B.12 Stochastic complements and minimum-norm deletion 135
B.13 Finite arithmetic checks of analytic formulas 136
B.14 Data limits 136

1 The problem and the results

For a real $n \times n$ matrix $A = (a_{ij})$, its permanent is

$$\operatorname{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Unlike the determinant, this sum has no permutation signs. A matrix is *doubly stochastic* if its entries are nonnegative and every row sum and column sum is one. Write Ω_n for this set, J_n for the matrix with every entry $1/n$, and

$$c_n = \frac{n!}{n^n}, \qquad n \geq 1.$$

The transpose is denoted by A^{T}. The expression A^2 denotes matrix multiplication, whereas $\operatorname{per}(A)^2$ is a scalar square.

1.1 Gyires's inequality

Gyires's conjecture [10] appears as Conjecture 20 in Cheon and Wanless [5]. For real doubly stochastic matrices, it asks for

$$\Delta_n(A) := 4\,\mathrm{per}(A)^2 - c_n\left[\mathrm{per}(AA^{\mathsf{T}}) + \mathrm{per}(A^{\mathsf{T}}A) + 2\,\mathrm{per}(A^2)\right] \geq 0, \tag{1}$$

with equality exactly at $A = J_n$. The corresponding ratio form is

$$\frac{4\,\mathrm{per}(A)^2}{\mathrm{per}(AA^{\mathsf{T}}) + \mathrm{per}(A^{\mathsf{T}}A) + 2\,\mathrm{per}(A^2)} \geq c_n.$$

Its denominator is positive. Indeed, each diagonal entry of AA^{T} is the positive squared norm of a row of A. The identity permutation contributes the product of these entries to $\mathrm{per}(AA^{\mathsf{T}})$. All remaining contributions are nonnegative.

Both sides of (1) have degree $2n$ in the entries. The homogeneous inequality therefore also makes sense without stochastic normalization. A sufficient stronger pair of bounds is

$$c_n\,\mathrm{per}(AA^{\mathsf{T}}) \leq \mathrm{per}(A)^2, \qquad c_n\,\mathrm{per}(A^{\mathsf{T}}A) \leq \mathrm{per}(A)^2. \tag{2}$$

Section 2 proves the permanent comparison that makes this implication precise.

1.2 Matrix results and proof methods

Sections 3 and 4 treat unrestricted nonnegative circulant weights in orders five, six, and seven. They give equality cases and sharp sorted-gap margins. The order-five, order-six, permutation-polynomial, four-block, and exceptional-vector theorems have formal proofs in Lean. The order-six and order-seven proofs use exhaustive exact coefficient certificates.

Section 5 treats $aI + bQ + cQ^2$ in every dimension, where Q is any permutation matrix and $a, b, c \geq 0$. The proof combines actual transfer identities, a weighted trace estimate from order 23, exact polynomial identities below that threshold, and a cycle decomposition. Lean checks the coefficient identities and the final matrix theorem. The common permutation matrix is essential to the cycle decomposition.

Section 6 treats doubly stochastic matrices with at most two distinct row vectors and two distinct column vectors. This gives four constant blocks, with unequal block sizes and boundary entries permitted. The proof gives the contractions $\mathrm{per}(AA^{\mathsf{T}}), \mathrm{per}(A^{\mathsf{T}}A), \mathrm{per}(A^2) \leq \mathrm{per}(A)$. Section 7 proves both strong bounds when all but one column are identical, or when all but one row are identical. The exceptional vector can have arbitrarily many distinct entries. Section 8 proves the contractions for every rank-at-most-two doubly stochastic matrix in orders 2 through 15. This finite-order result does not establish the all-order rank-two claim.

Section 9 proves closure properties and sparse-perturbation results. It distinguishes a one-sided bound for three exceptional rows in orders at least five from the two-sided theorem in orders at least 31. It also proves a fixed-order reduction to the matrix boundary.

1.3 Supporting results and their limits

The polynomial approach leads to scalar localization, mean-one and symmetric-reference norm inequalities, rank-two scalar reductions, deletion and balancing results, local polynomial inequalities, and exact expectations in a regular bipartite configuration model. These results isolate scalar and polynomial inequalities that would imply further matrix bounds. The

remaining inequalities are stated explicitly, separate from the proved reductions. Exact counterexamples rule out several stronger auxiliary statements. None contradicts Gyires's conjecture.

1.4 Verification status

The principal order-five and order-six inequalities, equality cases, and sharp margins have complete Lean proofs. The all-order permutation-polynomial, four-block, and exceptional-vector theorems also have complete Lean proofs. General permanent comparisons and direct-sum closure also have Lean proofs. The rank-two theorem has a complete Lean proof for orders 2 through 8. The order-seven circulant theorem and rank-two orders 9 through 15 remain incomplete in Lean. Section 17 maps the formalized theorems to their Lean declarations and explains how to reproduce the computations.

1.5 Prior work

The van der Waerden lower bound gives $\operatorname{per}(A) \geq c_n$ on Ω_n, with equality only at J_n. The classical proofs are recorded in [11, 7, 8]. The four-block argument below derives the needed lower bound directly for its own class.

Chang [4] proved the order-three Gyires inequality. His sufficient condition in every order requires each entry to be at least $(n-2)/(n-1)^2$. It does not include zero entries. Ozols [15] proves and formalizes the full order-four result and a stronger Gram-permanent bound.

Minc [14, Theorem 2] gave the classical permanent formula for three-term circulants. The permutation-polynomial section uses that formula and supplies its transfer derivation, including the small orders. The probability argument uses the classical fixed-mean extremal reduction of Hoeffding [12] and Newton's inequalities. Both results provide the basis for the four-block proof.

The order-five theorem and its formalization are also available in [9]. The literature comparison for the broader results is incomplete, so no priority claim is made.

2 Common permanent identities

The matrix arguments use the same comparison between a matrix square and two Gram matrices. This section proves the comparison before applying any structural restriction.

For a real matrix M, define

$$p_M(z) = \prod_{i=1}^{n} \left(\sum_{j=1}^{n} M_{ij} z_j \right).$$

For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$, write $|\alpha| = \sum_j \alpha_j$, $\alpha! = \prod_j \alpha_j!$, and $z^\alpha = \prod_j z_j^{\alpha_j}$. If $p = \sum_{|\alpha|=n} p_\alpha z^\alpha$ and $q = \sum_{|\alpha|=n} q_\alpha z^\alpha$, set

$$\langle p, q \rangle_F = \sum_{|\alpha|=n} \alpha! p_\alpha q_\alpha.$$

All weights $\alpha!$ are positive, so this defines an inner product on the degree-n coefficient vectors, called the Fischer inner product.

Lemma 2.1 (Permanent coefficient identity)**.** *For real $n \times n$ matrices M, N,*

$$\langle p_M, p_N \rangle_F = \operatorname{per}(MN^{\mathsf{T}}), \qquad [z_1 \cdots z_n] p_M = \operatorname{per}(M). \tag{3}$$

Proof. The second identity selects each column exactly once in the product of row forms. For the first identity, expand

$$\operatorname{per}(MN^{\mathsf{T}}) = \sum_{\sigma \in S_n} \sum_{\varphi:[n]\to[n]} \prod_i M_{i,\varphi(i)} \prod_j N_{j,\varphi(\sigma^{-1}(j))}.$$

Fix a map φ whose column multiplicities are α. As σ varies, each map with the same multiplicities occurs exactly $\alpha!$ times in $\varphi \circ \sigma^{-1}$. For each column, its occurrences can be matched in $\alpha_j!$ ways. Multiplication over columns gives $\alpha!$. Summing first over maps with multiplicity α, then over α, proves the identity. □

Lemma 2.2 (Permanent Cauchy–Schwarz)**.** *For real $n \times n$ matrices M, N,*

$$|\operatorname{per}(MN^{\mathsf{T}})| \le \sqrt{\operatorname{per}(MM^{\mathsf{T}}) \operatorname{per}(NN^{\mathsf{T}})}.$$

In particular, for a nonnegative matrix A,

$$\operatorname{per}(A^2) \le \sqrt{\operatorname{per}(AA^{\mathsf{T}}) \operatorname{per}(A^{\mathsf{T}}A)}.$$

Proof. Apply Cauchy–Schwarz to the inner product in Lemma 2.1. For the second statement, take $M = A$ and $N = A^{\mathsf{T}}$. □

Lemma 2.3. *For real $n \times n$ matrices M, N,*

$$2 \operatorname{per}(MN^{\mathsf{T}}) \le \operatorname{per}(MM^{\mathsf{T}}) + \operatorname{per}(NN^{\mathsf{T}}).$$

For a nonnegative matrix A, the two bounds (2) *imply $\Delta_n(A) \ge 0$.*

Proof. The coefficient identity and $2uv \le u^2 + v^2$ give the first statement term by term. For the second, put $x = \operatorname{per}(AA^{\mathsf{T}})$ and $y = \operatorname{per}(A^{\mathsf{T}}A)$. The first statement gives $2 \operatorname{per}(A^2) \le x + y$. Hence the bracket in (1) is at most $2(x+y)$. The two strong bounds give $c_n(x+y) \le 2 \operatorname{per}(A)^2$, as required. □

The proof applies to the original matrix indexing. No invariance of $\operatorname{per}(A^2)$ under independent row and column permutations is needed.

3 Nonnegative circulants of order five

The order-five inequality admits a sharp quantitative margin and an exact equality characterization. Its principal statements have complete Lean proofs. The proof covers all nonnegative weights, including zero and repeated weights. The certificate is finite, but it applies to every real point of the matrix family.

3.1 Notation

Indices of rows, columns and weights are taken in $\mathbb{Z}/5\mathbb{Z} = \{0,1,2,3,4\}$. For $w \in \mathbb{R}^5$, write $C(w)$ for the circulant with entries $C(w)_{ij} = w_{j-i}$, and

$$P(w) = \operatorname{per} C(w), \qquad g(w)_k = \sum_{i=0}^{4} w_i w_{i+k}, \qquad h(w)_k = \sum_{i=0}^{4} w_i w_{k-i} \qquad (k \in \mathbb{Z}/5\mathbb{Z}),$$

so that $g(w)$ is the circular autocorrelation of w and $h(w)$ its circular self-convolution. Note that $g(w)_{-k} = g(w)_k$. The *defects*, D_s for the strong inequality and D_G for Gyires's inequality, are the homogeneous polynomials of degree ten

$$D_s(w) = 625\,P(w)^2 - 24\,P(g(w)), \qquad D_G(w) = 625\,P(w)^2 - 12\big(P(g(w)) + P(h(w))\big). \quad (4)$$

Lemma 3.1. *For every $w \in \mathbb{R}^5$, $C(w)C(w)^{\mathsf{T}} = C(w)^{\mathsf{T}}C(w) = C(g(w))$ and $C(w)^2 = C(h(w))$.*

Proof. Write $C = C(w)$. Substituting $m = k - j$, respectively $m = i - k$ and $m = k - i$,

$$(CC^{\mathsf{T}})_{ij} = \sum_k w_{k-i}w_{k-j} = \sum_m w_m w_{m+j-i} = g_{j-i},$$
$$(C^{\mathsf{T}}C)_{ij} = \sum_k w_{i-k}w_{j-k} = \sum_m w_m w_{m+j-i} = g_{j-i},$$

and $(C^2)_{ij} = \sum_k w_{k-i}w_{j-k} = \sum_m w_m w_{(j-i)-m} = h_{j-i}$. □

Lemma 3.2. *For every $w = (w_0, \dots, w_4) \in \mathbb{R}^5$, writing $w = (a,b,c,d,e)$,*

$$\begin{aligned} P(w) = {} & a^5 + b^5 + c^5 + d^5 + e^5 + 15\,abcde \\ & + 5\big(a^3be + a^3cd + a^2b^2d + a^2bc^2 + a^2ce^2 + a^2d^2e + ab^3c + ab^2e^2 + abd^3 + ac^3e \\ & \quad + ac^2d^2 + ade^3 + b^3de + b^2c^2e + b^2cd^2 + bc^3d + bce^3 + bd^2e^2 + c^2de^2 + cd^3e\big). \end{aligned}$$

Proof. By definition $P(w) = \sum_{\sigma \in S_5} \prod_i w_{\sigma(i)-i}$, a sum of 120 monomials of degree five; collecting them gives the 26 terms displayed. For instance, the five translations $\sigma(i) = i + k$ give w_k^5, and the monomial $abcde$ arises from the 15 permutations σ for which $i \mapsto \sigma(i) - i$ is a bijection of $\mathbb{Z}/5\mathbb{Z}$. The coefficients add up to $120 = 5!$. □

Corollary 3.3. *For every $w \in \mathbb{R}^5$, $P(h(w)) \le P(g(w))$. Consequently $D_G(w) \ge D_s(w)$.*

Proof. Take $M = C(w)$ and $N = C(w)^{\mathsf{T}}$ in Lemma 2.3. By Lemma 3.1, $MN^{\mathsf{T}} = C(h(w))$ and $MM^{\mathsf{T}} = NN^{\mathsf{T}} = C(g(w))$, so $2P(h(w)) \le 2P(g(w))$. Finally $D_G - D_s = 12\big(P(g) - P(h)\big)$. □

3.2 The inequalities and sharp margin

Theorem 3.4. *Let $w \in \mathbb{R}^5$ with $w_i \ge 0$ for all i, and let $C = C(w)$. Then*

$$24\operatorname{per}(CC^{\mathsf{T}}) \le 625\operatorname{per}(C)^2, \qquad \textit{that is,} \qquad \frac{5!}{5^5}\operatorname{per}(CC^{\mathsf{T}}) \le \operatorname{per}(C)^2.$$

Equality holds if and only if $w_0 = w_1 = w_2 = w_3 = w_4$.

Theorem 3.5. *Let $w \in \mathbb{R}^5$ with $w_i \geq 0$ for all i, let $C = C(w)$, let π be a permutation of $\mathbb{Z}/5\mathbb{Z}$ with $w_{\pi(0)} \leq w_{\pi(1)} \leq \cdots \leq w_{\pi(4)}$, and put $z_j = w_{\pi(j)} - w_{\pi(j-1)}$ for $1 \leq j \leq 4$. Then*

$$D_s(w) = 625\,\mathrm{per}(C)^2 - 24\,\mathrm{per}(CC^{\mathsf{T}}) \;\geq\; 484\,(z_1^{10} + z_2^{10} + z_3^{10} + z_4^{10}).$$

Equality holds for $w = (1, 1, 0, 0, 0)$, for which $z_1^{10} + \cdots + z_4^{10} = 1$ and $D_s(w) = 484$. Consequently, if a real constant K satisfies $D_s(w) \geq K\,(z_1^{10} + \cdots + z_4^{10})$ for all such w and π, then $K \leq 484$.

Corollary 3.6. *Let $w \in \mathbb{R}^5$ with $w_i \geq 0$ for all i, and let $C = C(w)$. Then*

$$24\big(\mathrm{per}(CC^{\mathsf{T}}) + \mathrm{per}(C^{\mathsf{T}}C) + 2\,\mathrm{per}(C^2)\big) \leq 2500\,\mathrm{per}(C)^2,$$

with equality if and only if all five weights are equal. If moreover $w_0 + \cdots + w_4 = 1$, then $C \in \Omega_5$, the denominator below is positive, and

$$\frac{4\,\mathrm{per}(C)^2}{\mathrm{per}(CC^{\mathsf{T}}) + \mathrm{per}(C^{\mathsf{T}}C) + 2\,\mathrm{per}(C^2)} \;\geq\; \frac{24}{625} = \frac{5!}{5^5},$$

with equality if and only if $C = J_5$. Thus Gyires's inequality (1) *holds for every doubly stochastic circulant matrix of order five.*

3.3 Certificate method

The idea in brief. By Corollary 3.3 it suffices to bound D_s. The defect is invariant under the affine changes of indices $i \mapsto ai + b$ of $\mathbb{Z}/5\mathbb{Z}$, and these reduce the 120 possible orderings of the weights to six. In each of the six chambers, the defect written in the smallest weight and the gaps between consecutive sorted weights is a polynomial with nonnegative integer coefficients, which proves the inequality, the margin and the equality case at once.

3.4 Affine symmetry

For $a \in (\mathbb{Z}/5\mathbb{Z})^\times$ and $b \in \mathbb{Z}/5\mathbb{Z}$ let $(T_{a,b}w)_i = w_{ai+b}$; the twenty maps $i \mapsto ai + b$ are the affine bijections of $\mathbb{Z}/5\mathbb{Z}$.

Lemma 3.7. *For all $w \in \mathbb{R}^5$, a and b: $P(T_{a,b}w) = P(w)$, $g(T_{a,b}w)_k = g(w)_{ak}$ and $h(T_{a,b}w)_k = h(w)_{ak+2b}$. Consequently $D_s(T_{a,b}w) = D_s(w)$ and $D_G(T_{a,b}w) = D_G(w)$.*

Proof. The identity $C(T_{a,b}w)_{ij} = w_{a(j-i)+b} = C(w)_{ai,\,aj+b}$ shows that $C(T_{a,b}w)$ arises from $C(w)$ by permuting the rows by $i \mapsto ai$ and the columns by $j \mapsto aj + b$; permanents are invariant under such permutations. Substituting $m = ai + b$,

$$g(T_{a,b}w)_k = \sum_i w_{ai+b}\,w_{a(i+k)+b} = \sum_m w_m\,w_{m+ak},$$
$$h(T_{a,b}w)_k = \sum_i w_{ai+b}\,w_{a(k-i)+b} = \sum_m w_m\,w_{ak+2b-m}.$$

Hence $g(T_{a,b}w) = T_{a,0}\,g(w)$ and $h(T_{a,b}w) = T_{a,2b}\,h(w)$, and the first part gives $P(g(T_{a,b}w)) = P(g(w))$ and $P(h(T_{a,b}w)) = P(h(w))$. □

3.5 Six chambers

A *canonical ordering* is one of the six sequences

$$01234, \quad 01243, \quad 01324, \quad 01342, \quad 01423, \quad 01432, \tag{5}$$

that is, an ordering $\kappa = (\kappa(0), \dots, \kappa(4))$ of $\mathbb{Z}/5\mathbb{Z}$ with $\kappa(0) = 0$ and $\kappa(1) = 1$. For a canonical ordering κ and $z = (z_0, \dots, z_4) \in \mathbb{R}^5$ let $L_\kappa(z) \in \mathbb{R}^5$ be the vector with

$$L_\kappa(z)_{\kappa(j)} = z_0 + z_1 + \cdots + z_j \qquad (0 \le j \le 4).$$

If $z \ge 0$, the weights of $L_\kappa(z)$ are nondecreasing along κ, z_0 is the smallest weight and $z_1, \dots, z_4$ are the gaps between consecutive sorted weights. The *chamber* of κ is the set $\{L_\kappa(z) : z \ge 0\}$ of nonnegative weight vectors that are nondecreasing along κ.

Lemma 3.8. *Let π be a permutation of $\mathbb{Z}/5\mathbb{Z}$. There are $a \in (\mathbb{Z}/5\mathbb{Z})^\times$, $b \in \mathbb{Z}/5\mathbb{Z}$ and a canonical ordering κ with $\pi(j) = a\kappa(j) + b$ for every j, and the triple (a, b, κ) is unique. If $w \in \mathbb{R}^5$ satisfies $w_{\pi(0)} \le \cdots \le w_{\pi(4)}$ and $w_{\pi(0)} \ge 0$, then $T_{a,b}w = L_\kappa(z)$ with $z_0 = w_{\pi(0)}$ and $z_j = w_{\pi(j)} - w_{\pi(j-1)}$ $(1 \le j \le 4)$, and $z \ge 0$.*

Proof. Put $b = \pi(0)$ and $a = \pi(1) - \pi(0) \ne 0$, and $\kappa(j) = a^{-1}(\pi(j) - b)$. Then κ is an ordering of $\mathbb{Z}/5\mathbb{Z}$ with $\kappa(0) = 0$ and $\kappa(1) = 1$, hence canonical, and $\pi(j) = a\kappa(j) + b$; conversely the conditions $\kappa(0) = 0$ and $\kappa(1) = 1$ force $b = \pi(0)$ and $a = \pi(1) - \pi(0)$. For the second statement, $(T_{a,b}w)_{\kappa(j)} = w_{a\kappa(j)+b} = w_{\pi(j)} = z_0 + \cdots + z_j$, and $z \ge 0$ because the sorted weights are nondecreasing and the smallest is nonnegative. □

3.6 The chamber certificate

Lemma 3.9. *For each canonical ordering κ, the polynomial*

$$R_\kappa(z) = D_s(L_\kappa(z)) - 484\,(z_1^{10} + z_2^{10} + z_3^{10} + z_4^{10})$$

in $z = (z_0, \dots, z_4)$ has nonnegative integer coefficients. More precisely, $D_s(L_\kappa(z))$ is homogeneous of degree ten, so it has $\binom{14}{4} = 1001$ coefficients; 996 of them are positive, and the five coefficients of z_0^{10} and $z_0^9 z_j$ $(1 \le j \le 4)$ are zero. The coefficients of $z_1^{10}, z_2^{10}, z_3^{10}, z_4^{10}$ are 233 296, 44 209, 484 and 601, and every positive coefficient is at least 484.

Proof. This is a finite computation with integer polynomials. By Lemma 3.2, $P(w)$ is an explicit integer polynomial, and by definition so is $P(g(w))$; one substitutes $w = L_\kappa(z)$ into $625P(w)^2 - 24P(g(w))$, expands, and reads off the 1001 coefficients, for each of the six orderings (5). All coefficient claims of the lemma are obtained in this way. The statement about R_κ follows: the four pure powers z_j^{10}, $j \ge 1$, have coefficients at least 484, and every other coefficient is unchanged. □

The Lean proof organizes this computation by powers of the smallest weight and the first gap (Section 17). Let $u = L_\kappa(0, z_1, \dots, z_4)$, so that $L_\kappa(z) = z_0\mathbf{1} + u$, where $\mathbf{1} = (1, \dots, 1)$. Two identities make the dependence on z_0 explicit. First, $P(t\mathbf{1} + u) = \sum_{k=0}^{5} p_k(u)\, t^k$ for explicit polynomials p_k. Second, $g(t\mathbf{1} + u) = (5t^2 + 2tS)\mathbf{1} + g(u)$, where $S = u_0 + \cdots + u_4$, so that $P(g(t\mathbf{1} + u)) = \sum_{k=0}^{5} p_k(g(u))\,(5t^2 + 2tS)^k$. Hence $D_s(L_\kappa(z)) = \sum_{j=0}^{10} Q_j\, z_0^j$ with explicit polynomials Q_j in $z_1, \dots, z_4$, and each Q_j is expanded in powers of z_1 with coefficients that are polynomials in z_2, z_3, z_4; by Lemma 3.9, $Q_9 = Q_{10} = 0$. The nonnegativity of the coefficients of R_κ amounts to the statement that, after $484\,(z_1^{10} + \cdots + z_4^{10})$ is subtracted from Q_0, each of these polynomials in z_2, z_3, z_4 has nonnegative coefficients.

3.7 Proof of Theorems 3.5 and 3.4

Proof of Theorem 3.5. Let $w \geq 0$ and π be as in the theorem, and choose a, b, κ and z by Lemma 3.8. By Lemmas 3.7 and 3.9, since $z \geq 0$,

$$D_s(w) = D_s(T_{a,b}w) = D_s(L_\kappa(z)) = 484\,(z_1^{10} + \cdots + z_4^{10}) + R_\kappa(z) \;\geq\; 484\,(z_1^{10} + \cdots + z_4^{10}).$$

For the equality case, let $w = (1, 1, 0, 0, 0)$, sorted by $\pi = (2, 3, 4, 0, 1)$, with gaps $(z_1, z_2, z_3, z_4) = (0, 0, 1, 0)$. Here $g(w) = (2, 1, 0, 0, 1)$, and Lemma 3.2 gives $P(w) = 2$ and $P(g(w)) = 84$, so $D_s(w) = 625{\cdot}4 - 24{\cdot}84 = 484$. A constant K as in the theorem therefore satisfies $K{\cdot}1 \leq 484$. □

Proof of Theorem 3.4. By Lemma 3.1, $\operatorname{per}(CC^\mathsf{T}) = P(g(w))$, so the inequality is $D_s(w) \geq 0$, which follows from Theorem 3.5 applied with any permutation π that sorts w. If $D_s(w) = 0$, Theorem 3.5 gives $z_1 = \cdots = z_4 = 0$, so all weights are equal. Conversely, if every weight equals c, then $P(w) = 120c^5$ (each of the 120 permutations contributes c^5) and $g(w) = (5c^2, \ldots, 5c^2)$, so $P(g(w)) = 120 \cdot 5^5 c^{10}$ and $625P(w)^2 = 9\,000\,000\,c^{10} = 24P(g(w))$. □

3.8 Proof of Corollary 3.6

Proof. By Lemma 3.1, $\operatorname{per}(CC^\mathsf{T}) + \operatorname{per}(C^\mathsf{T}C) + 2\operatorname{per}(C^2) = 2P(g(w)) + 2P(h(w))$, so the first inequality reads $4D_G(w) \geq 0$. By Corollary 3.3 and Theorem 3.4, $D_G(w) \geq D_s(w) \geq 0$. If $D_G(w) = 0$, then $D_s(w) = 0$, and all weights are equal by Theorem 3.4; conversely, for constant w one has $h(w) = g(w)$, hence $D_G(w) = D_s(w) = 0$.

Now let also $\sum_i w_i = 1$. Every row and every column of C contains each weight once, so $C \in \Omega_5$. Some weight is positive, so $g(w)_0 = \sum_i w_i^2 > 0$. The matrix $C(g(w))$ has nonnegative entries and diagonal entries $g(w)_0$, so $P(g(w)) \geq g(w)_0^5 > 0$, the identity permutation contributing $g(w)_0^5$ and the others nonnegative terms. Since also $P(h(w)) \geq 0$, the denominator $2P(g(w)) + 2P(h(w))$ is positive, and dividing the first inequality by 625 times the denominator gives the ratio inequality. Equality holds if and only if all weights are equal, that is, equal to $1/5$, which means $C = J_5$. □

4 Nonnegative circulants of orders six and seven

The sorted-gap method in Section 3 also gives complete finite certificates in orders six and seven. The certificates give sharp margins for every nonnegative real weight vector in these two orders.

For this section, let indices lie in $\mathbb{Z}/n\mathbb{Z}$, and define

$$C_n(w)_{ij} = w_{j-i}, \qquad g_j(w) = \sum_{i=0}^{n-1} w_i w_{i+j}, \qquad \mathcal{P}_n(w) = \operatorname{per}(C_n(w)).$$

Direct multiplication gives

$$C_n(w)C_n(w)^\mathsf{T} = C_n(g(w)) = C_n(w)^\mathsf{T}C_n(w).$$

In particular, every real circulant is normal. Define its homogeneous strong deficit by

$$D_n(w) = n^n \mathcal{P}_n(w)^2 - n!\mathcal{P}_n(g(w)). \tag{6}$$

This is an integer polynomial of degree $2n$.

Choose any permutation π that sorts the weights, and set

$$z_0 = w_{\pi(0)}, \qquad z_j = w_{\pi(j)} - w_{\pi(j-1)} \quad (1 \le j \le n-1).$$

For nonnegative weights, all these coordinates are nonnegative. The positive-index coordinates are the consecutive sorted gaps.

Theorem 4.1 (Sharp margins in orders six and seven)**.** *For every nonnegative circulant $C = C_6(w)$,*

$$324\,\mathrm{per}(C)^2 - 5\,\mathrm{per}(CC^{\mathsf{T}}) \ge 296 \sum_{j=1}^{5} z_j^{12}. \tag{7}$$

For every nonnegative circulant $C = C_7(w)$,

$$117649\,\mathrm{per}(C)^2 - 720\,\mathrm{per}(CC^{\mathsf{T}}) \ge 116929 \sum_{j=1}^{6} z_j^{14}. \tag{8}$$

Both constants on the right are optimal for inequalities of these forms. For each of these two orders,

$$c_n\,\mathrm{per}(CC^{\mathsf{T}}) \le \mathrm{per}(C)^2,$$

with equality if and only if all weights are equal. If the weights sum to one, then C satisfies Gyires's inequality, with equality if and only if $C = J_n$.

The theorem includes zero weights and equal weights. The zero vector is included in the homogeneous equality case. The sharpness assertion does not classify all equality cases of the margin inequalities themselves.

4.1 Affine symmetry and complete coverage

Let a be a unit modulo n, and let $b \in \mathbb{Z}/n\mathbb{Z}$. Write $(T_{a,b}w)_j = w_{aj+b}$. Then

$$C_n(T_{a,b}w)_{ij} = C_n(w)_{ai,aj+b}.$$

This is a row permutation and a column permutation of $C_n(w)$. Both $\mathrm{per}(C_n(w))$ and $\mathrm{per}(C_n(w)C_n(w)^{\mathsf{T}})$ are unchanged. Thus

$$D_n(T_{a,b}w) = D_n(w). \tag{9}$$

The argument concerns the Gram product. It does not assume that independent row and column permutations preserve a matrix square.

The affine group acts freely on the ordered lists of all n indices. Indeed, an affine map that fixes such a list fixes every residue. It therefore has $a = 1$ and $b = 0$. Each orbit has $n\varphi(n)$ elements, where $\varphi(n)$ counts the units modulo n. Each orbit has one selected sorting representative. To construct it, first translate the first index to zero. Then take the lexicographically least list among all its unit multiples. This gives

$$\frac{6!}{6\varphi(6)} = 60 \quad \text{and} \quad \frac{7!}{7\varphi(7)} = 120$$

representatives in orders six and seven, respectively.

This construction matters in composite order. In order six, a difference between two distinct indices need not be a unit. One cannot always normalize the first two indices to 0 and 1, as

in order five. Translation and minimization over unit multiples avoid that restriction. The checking program also constructs every affine image of every representative. It verifies coverage of all 720 and 5040 sorting permutations in the two orders.

For each representative π, substitute

$$w_{\pi(j)} = z_0 + \cdots + z_j \qquad (0 \le j \le n-1). \tag{10}$$

Every nonnegative real weight vector belongs to one of the resulting closed sorting regions, after an affine relabeling. Ties cause no omission. Choose any sorting permutation when weights are equal; the corresponding gaps can be zero.

4.2 Construction and independent checking of the polynomials

The primary program reconstructs each permanent by its defining subset recurrence. For a matrix $M = (M_{ij})$ with polynomial entries, let $H_\varnothing = 1$. For every nonempty subset S of columns, set

$$H_S = \sum_{j \in S} H_{S \setminus \{j\}} M_{|S|-1,j}. \tag{11}$$

After all rows are processed, $H_{\{0,\ldots,n-1\}} = \operatorname{per}(M)$. Each recurrence term selects one unused column for the next row. Every permutation is selected once. The program applies this recurrence to $C_n(w)$ and $C_n(g(w))$, then forms (6).

Polynomial coefficients are arbitrary-precision Python integers. The monomial with exponent vector $(\alpha_0, \ldots, \alpha_{n-1})$ is encoded by

$$\sum_{j=0}^{n-1} \alpha_j (2n+1)^j.$$

Every intermediate exponent is at most $2n$. Thus addition of encoded exponents during polynomial multiplication produces no carry between positions. The encoding preserves the full multivariate polynomial.

To implement (10), the program first relabels the variables by π. It then makes the elementary substitutions $z_j \leftarrow z_j + z_{j-1}$, in the order $j = n-1, n-2, \ldots, 1$. The binomial theorem gives every new coefficient exactly. The descending order turns each original variable of sorted rank j into $z_0 + \cdots + z_j$.

A separate checker reconstructs the original permanent by enumeration of all permutations. For the Gram permanent, it groups permutations by their autocorrelation displacement counts and uses $g_j = g_{n-j}$. It compares the resulting original deficit polynomial with the primary polynomial, coefficient by coefficient. Its monomial encoding uses fixed-width bit positions with base greater than $2n$.

The checker then verifies every transformed polynomial by inverse binomial substitutions. It replaces z_j by $z_j - z_{j-1}$ in the order $j = 1, 2, \ldots, n-1$. For the two-variable part of fixed total degree t, if the input coefficient of $z_j^e z_{j-1}^{t-e}$ is a_e, the inverse coefficient of $z_j^h z_{j-1}^{t-h}$ is

$$\sum_{e=h}^{t} (-1)^{e-h} \binom{e}{h} a_e. \tag{12}$$

After all inverse substitutions, the result must equal the independently reconstructed deficit with its variables permuted by π. This verifies every full sorting-region polynomial, not only selected values of the weights.

Lemma 4.2 (Complete coefficient certificate)**.** *For every affine sorting representative in order six or seven, the polynomial obtained from* (6) *by* (10) *has nonnegative integer coefficients. Its only zero coefficient positions are*

$$z_0^{2n}, \qquad z_0^{2n-1} z_j \quad (1 \le j \le n-1).$$

The exact coefficient data are

n	*Regions*	*Positions per region*	*Positive*	*Zero*	*Least positive*
6	*60*	*6188*	*6182*	*6*	*42624*
7	*120*	*38760*	*38753*	*7*	*818503*

The last column gives the minimum over all regions of the stated order.

Exact computational proof. The primary generator and the independent checker execute the finite constructions above. The checker verifies affine coverage, equality of the original polynomials, and every inverse identity. It also checks that every stored coefficient is positive. There are $\binom{3n-1}{n-1}$ possible monomials of degree $2n$ in n variables. In every region, the checker finds exactly $\binom{3n-1}{n-1} - n$ positive terms and verifies that none occupies one of the n positions displayed in the lemma. The remaining positions are therefore exactly those zeros.

The zero positions also follow directly from the polynomial. At $w = t\mathbf{1}$,

$$\mathcal{P}_n(w) = n! t^n, \qquad \mathcal{P}_n(g(w)) = n! n^n t^{2n}.$$

Both D_n and each of its first partial derivatives vanish there. For $t > 0$, the derivatives of these two permanent polynomials with respect to any weight are $n! t^{n-1}$ and $2\, n!\, n^n t^{2n-1}$, respectively. These identities give the zero coefficients of z_0^{2n} and $z_0^{2n-1} z_j$ after the gap substitution. They are polynomial identities, so the zero baseline causes no exception.

The calculation checks $60 \cdot 6188 + 120 \cdot 38760 = 5\,022\,480$ coefficient positions. The exact output records include each region's positive count, minimum, pure-gap coefficients, and direct integer examples. The programs reconstruct the full polynomials; the JSON records are summaries and do not store all coefficient arrays. □

4.3 Sharpness, equality, and the Gyires deduction

Proof of Theorem 4.1. Let $M_6 = 42624$ and $M_7 = 818503$. Lemma 4.2 shows that every pure gap term z_j^{2n}, $j \ge 1$, has coefficient at least M_n. Subtracting $M_n \sum_{j=1}^{n-1} z_j^{2n}$ therefore leaves a polynomial with nonnegative coefficients in each sorting region. Equations (9) and (10) then give, for every nonnegative weight vector,

$$D_n(w) \ge M_n \sum_{j=1}^{n-1} z_j^{2n}. \tag{13}$$

Dividing by $144 = \gcd(6^6, 6!)$ gives (7). Dividing by $7 = \gcd(7^7, 7!)$ gives (8).

For order six, take $w = (1, 1, 0, 0, 0, 0)$. Direct permanent evaluation gives $\operatorname{per}(C) = 2$ and $\operatorname{per}(CC^{\mathsf{T}}) = 200$. There is one gap equal to one, and all other gaps are zero. The reduced deficit is

$$324 \cdot 2^2 - 5 \cdot 200 = 296.$$

No larger universal coefficient can replace 296. For order seven, take $w = (1, 0, 0, 0, 0, 0, 0)$. Here $C = I$, both permanents equal one, and the gap sum is one. The reduced deficit is $117649 - 720 = 116929$. This proves the second sharpness assertion.

If the strong deficit vanishes, (13) forces all consecutive gaps to vanish. Thus all weights are equal. Conversely, direct evaluation at constant weights gives zero deficit.

Normality and Lemma 2.2 give $\operatorname{per}(C^2) \le \operatorname{per}(CC^{\mathsf{T}}) = \operatorname{per}(C^{\mathsf{T}}C)$. Consequently,

$$\Delta_n(C) \ge 4[\operatorname{per}(C)^2 - c_n \operatorname{per}(CC^{\mathsf{T}})] \ge \frac{4M_n}{n^n} \sum_{j=1}^{n-1} z_j^{2n}. \tag{14}$$

When the weights sum to one, C is doubly stochastic. If it differs from J_n, at least one sorted gap is positive, so Gyires's inequality is strict. At J_n, all three product matrices are J_n, and equality holds. □

4.4 A consequence in unbounded dimension

The fixed-order results also imply a support-restricted family in unbounded dimension. This consequence does not establish the strong inequality for unrestricted circulants of higher order.

Corollary 4.3 (Subgroup-coset support)**.** *Let $m \in \{5, 6, 7\}$, let $k \ge 1$, and set $n = km$. Suppose the nonnegative weights of an order-n circulant are supported in one coset of the subgroup of $\mathbb{Z}/n\mathbb{Z}$ of size m. Then*

$$c_n \operatorname{per}(CC^{\mathsf{T}}) \le \operatorname{per}(C)^2.$$

If $k > 1$ and $C \neq 0$, this inequality is strict. After stochastic normalization, Gyires's inequality also holds and is strict when $k > 1$.

Proof. A cyclic shift of the weights moves the support to the subgroup itself. This is a column permutation of C, which preserves the permanent and the Gram matrix. Group the row and column indices by their residues modulo k. The shifted matrix is then a direct sum of k identical $m \times m$ circulants B. Theorem 3.4 for $m = 5$ and Theorem 4.1 for $m = 6, 7$ give the strong inequality for each block. If the common row sum is zero, $C = 0$ and the claim is immediate. Otherwise divide every block by that positive row sum. The resulting blocks are doubly stochastic. Lemma 9.1 applies to their direct sum. Multiplying back and using degree-$2n$ homogeneity gives the strong inequality in dimension km. If $C \neq 0$, every identical block is nonzero. Its Gram permanent is positive because every Gram diagonal entry is positive. For $k > 1$, the strict dimension-constant comparison in that lemma therefore makes the strong inequality strict. The original matrix C is normal. The same permanent comparison as in (14) gives Gyires's inequality and its stated strictness. A nonzero nonnegative circulant has a positive common row sum, so stochastic normalization is available. □

4.5 Reproduction and limits

The certificate programs are `circ_gap_test.py` and `independent_certificate_check.py`. From a working directory containing both files, run

```
python3 circ_gap_test.py 6 all
python3 circ_gap_test.py 7 all
python3 independent_certificate_check.py 6
python3 independent_certificate_check.py 7
```

Run the independent checker without Python's `-O` option. It rejects optimized execution because that option disables assertions. Both programs return a nonzero exit code if a mathematical check fails or the required region coverage is incomplete. A complete check must report all 60 or 120 regions, as appropriate. The supporting source record gives the file hashes and regenerated outputs. Timing fields can differ between runs.

The exact programs use only the Python standard library. These programs give finite algebraic checks with complete parameter coverage in the two stated orders. The separate order-six Lean proof checks the actual matrix theorem, all sorting regions, both sharp margins, and their equality cases. The order-seven coefficient proof is not yet formalized in Lean. Selected-region experiments in orders eight and nine do not complete those orders and are not used in this theorem.

5 Permutation polynomials in arbitrary dimensions

The next theorem holds for every permutation cycle structure and all nonnegative weights. Its proof combines a weighted trace estimate with exact polynomial certificates.

Theorem 5.1. *Let $n \geq 1$, let Q be an $n \times n$ permutation matrix, and let $a, b, c \geq 0$. Put*

$$A = aI + bQ + cQ^2.$$

Then

$$c_n \operatorname{per}(AA^{\mathsf{T}}) \leq \operatorname{per}(A)^2, \qquad c_n \operatorname{per}(A^{\mathsf{T}}A) \leq \operatorname{per}(A)^2, \qquad c_n \operatorname{per}(A^2) \leq \operatorname{per}(A)^2. \tag{15}$$

The same hypotheses give

$$c_n [\operatorname{per}(AA^{\mathsf{T}}) + \operatorname{per}(A^{\mathsf{T}}A) + 2 \operatorname{per}(A^2)] \leq 4 \operatorname{per}(A)^2. \tag{16}$$

If $a + b + c = 1$, then A is doubly stochastic.

Set

$$C = AQ^{-1} = aQ^{-1} + bI + cQ. \tag{17}$$

A column permutation preserves the permanent, and $CC^{\mathsf{T}} = AA^{\mathsf{T}}$. It therefore suffices to prove the first bound for C. The proof first treats one cycle. It then combines the cycle bounds with the required dimension constants.

5.1 The cyclic permanent and its transfer matrix

Let P_n be the cyclic permutation matrix, and put $C_n = aP_n^{-1} + bI + cP_n$. Define

$$\lambda = \frac{b + \sqrt{b^2 + 4ac}}{2}, \qquad \nu = \frac{b - \sqrt{b^2 + 4ac}}{2}, \qquad L = \max\{a, c, \lambda\}. \tag{18}$$

The following formula is classical for three-term circulants. It appears in Minc's Theorem 2 [14]. The transfer argument also treats the coincident diagonals in orders one and two.

Lemma 5.2. *For every $n \geq 1$,*

$$\operatorname{per}(C_n) = a^n + c^n + \lambda^n + \nu^n \geq L^n. \tag{19}$$

Proof. The following transfer construction also computes the Gram permanent. For offsets $-r, \dots, r$, let $w_0, \dots, w_{2r}$ be the corresponding weights. Before processing row i, record the occupied columns $i-r, \dots, i+r-1$ in a mask m with $2r$ bits. A bit is one precisely when an earlier row has already selected that column. The new column $i+r$ has bit $2r$, initially zero.

Choose an unset bit $j \in \{0, \dots, 2r\}$ with weight w_j. Set that bit, and require the resulting bit zero to be one. The latter condition ensures that column $i-r$ is occupied before it leaves the window. The next mask is

$$m' = \left\lfloor \frac{m \text{ OR } 2^j}{2} \right\rfloor. \tag{20}$$

Here OR denotes bitwise union. The transfer matrix T sums the weights of these transitions.

To justify a cyclic boundary, lift the row and column indices to $\mathbb{Z}$. A cyclic matching with a chosen offset for each row lifts to a periodic matching. Its occupancy masks give a closed walk of length n in the transfer matrix. Conversely, each closed walk selects one column per row. The unset-bit rule prevents two selections of the same column. The outgoing-bit rule ensures that every column is selected. Periodicity then gives a matching of the cyclic indices. The states are determined by these choices, since each recorded bit leaves the window after finitely many steps. The two constructions preserve the product of the weights. Consequently,

$$\operatorname{per}\left(\sum_{j=0}^{2r} w_j P_n^{j-r}\right) = \operatorname{tr}(T^n). \tag{21}$$

If offsets coincide modulo n, the different offset choices contribute the corresponding entry multiplicity. Thus the identity remains valid for $n \le 2r$.

For C_n, use $r = 1$ and weights (a, b, c). Each transition adds one occupied bit and removes one occupied bit. The number of occupied bits is therefore constant. The four states split into scalar blocks a, c and the block

$$\begin{pmatrix} b & c \\ a & 0 \end{pmatrix}.$$

The latter block has eigenvalues λ, ν. The trace identity proves the equality in (19).

The all-forward and all-backward permutations give $\operatorname{per}(C_n) \ge a^n$ and $\operatorname{per}(C_n) \ge c^n$. To obtain the remaining lower bound, note that $\nu \le 0$ and $|\nu| \le \sqrt{ac}$. For even n, all four terms in the formula are nonnegative. For odd n,

$$a^n + c^n \ge 2(ac)^{n/2} \ge |\nu|^n.$$

Hence $\operatorname{per}(C_n) \ge \lambda^n$ in both cases. □

Set

$$d = a^2 + b^2 + c^2, \qquad e = b(a+c), \qquad f = ac. \tag{22}$$

The Gram matrix has the cyclic form

$$C_n C_n^{\mathsf{T}} = f P_n^{-2} + e P_n^{-1} + dI + e P_n + f P_n^2.$$

Its transfer matrix has sixteen states. The five population sizes give blocks of dimensions $1, 4, 6, 4, 1$. Both scalar blocks are f. In the mask orders $(1, 2, 4, 8)$ and $(7, 11, 13, 14)$, the two four-dimensional blocks are transposes. Their representative is

$$T_1(d, e, f) = \begin{pmatrix} e & d & e & f \\ f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \end{pmatrix}. \tag{23}$$

The six-dimensional block uses the mask order $(3, 5, 6, 9, 10, 12)$ and is

$$T_2(d, e, f) = \begin{pmatrix} d & e & 0 & f & 0 & 0 \\ e & 0 & e & 0 & f & 0 \\ f & 0 & 0 & 0 & 0 & 0 \\ 0 & e & d & 0 & 0 & f \\ 0 & f & 0 & 0 & 0 & 0 \\ 0 & 0 & f & 0 & 0 & 0 \end{pmatrix}. \tag{24}$$

Equation (21) gives

$$\operatorname{per}(C_n C_n^{\mathsf{T}}) = \operatorname{tr}(T_2^n) + 2\operatorname{tr}(T_1^n) + 2f^n. \tag{25}$$

The entries in these two matrices have degree two in a, b, c. They increase when any of d, e, f increases.

5.2 A uniform spectral bound

Write $\rho(M)$ for the spectral radius of a square matrix M. Set

$$\kappa = 1 + \sqrt{2}, \qquad \kappa^2 = 2\kappa + 1. \tag{26}$$

Lemma 5.3. *With the quantities in* (18) *and* (22)*,*

$$\rho(T_1(d, e, f)) \le \kappa L^2, \qquad \rho(T_2(d, e, f)) \le \kappa L^2.$$

The constant κ in the second bound is sharp.

Proof. If $L = 0$, all three weights are zero and both matrices vanish. Suppose $L > 0$ and divide the weights by L. It suffices to prove the result for $L = 1$. Then $0 \le a, c \le 1$ and $\lambda \le 1$. Since λ is the nonnegative root of $t^2 - bt - ac$,

$$0 \le b \le 1 - ac.$$

Increasing b to $1 - ac$ increases d and e and leaves f fixed. The spectral radius is monotone for nonnegative matrices under entrywise comparison. It is therefore enough to use $b = 1 - ac$.

Put $u = ac \in [0, 1]$. The inequality $(1 - a)(1 - c) \ge 0$ gives $a + c \le 1 + u$. Thus

$$\begin{aligned} d &= (a + c)^2 - 2u + (1 - u)^2 \le 2 - 2u + 2u^2 =: D(u), \\ e &= (1 - u)(a + c) \le 1 - u^2 =: E(u), \\ f &= u. \end{aligned} \tag{27}$$

It remains to bound $T_j(D(u), E(u), u)$ for $0 \le u \le 1$. In this interval, $D(u) \le 2$ and $E(u) \le 1$.

For T_1, use the positive column vector

$$w = (1, \kappa^{-1}, \kappa^{-2}, \kappa^{-3})^{\mathsf{T}}.$$

The last three coordinates of $T_1 w$ are at most those of κw. The first coordinate is at most

$$1 + \frac{2}{\kappa} + \frac{1}{\kappa^2} + \frac{1}{\kappa^3} = \frac{5}{\kappa} < \kappa.$$

Here the equality uses $\kappa^2 = 2\kappa + 1$, and the inequality uses $\kappa^2 > 5$. Thus $T_1 w \le \kappa w$. For a nonnegative matrix M and a positive vector w, $Mw \le \kappa w$ implies $\rho(M) \le \kappa$. Indeed,

conjugation by the diagonal matrix with diagonal w gives row sums at most κ. This proves the four-dimensional bound.

For the six-dimensional bound, first take $0 < u < 1$. Abbreviate $D = D(u)$ and $E = E(u)$, and define

$$\begin{aligned} v_1 &= 1, & v_2 &= \frac{E}{\kappa - u}, & v_3 &= \frac{u}{\kappa}, \\ v_4 &= \frac{E^2}{\kappa(\kappa - u)} + \frac{Du}{\kappa^2} + \frac{u^3}{\kappa^3}, & v_5 &= \frac{uE}{\kappa(\kappa - u)}, & v_6 &= \frac{u^2}{\kappa^2}. \end{aligned} \tag{28}$$

Every coordinate is positive. Substitution into (24) gives

$$(T_2 v)_i = \kappa v_i \qquad (2 \le i \le 6).$$

For example, the second coordinate follows from

$$E\left(1 + \frac{u}{\kappa}\right) + \frac{u^2 E}{\kappa(\kappa - u)} = \frac{\kappa E}{\kappa - u}.$$

The first coordinate has residual

$$\begin{aligned} \kappa v_1 - (T_2 v)_1 &= \kappa - D\left(1 + \frac{u^2}{\kappa^2}\right) - \frac{E^2(\kappa + u)}{\kappa(\kappa - u)} - \frac{u^4}{\kappa^3} \\ &= \frac{P(u)}{\kappa^3(\kappa - u)}, \end{aligned} \tag{29}$$

where

$$\begin{aligned} P(u) = \kappa^4(\kappa - u) - D\kappa(\kappa - u)(\kappa^2 + u^2) \\ - E^2\kappa^2(\kappa + u) - u^4(\kappa - u). \end{aligned}$$

Expansion with $\kappa^2 = 2\kappa + 1$ gives the exact factorization

$$P(u) = u(1 - u)\big[(20\kappa + 8) - (8\kappa + 4)u + (12\kappa + 4)u^2\big]. \tag{30}$$

The bracket equals

$$(12\kappa + 4) + (8\kappa + 4)(1 - u) + (12\kappa + 4)u^2,$$

which is positive on $[0, 1]$. The denominator in (29) is positive. Hence $T_2 v \le \kappa v$, and the positive-vector comparison gives $\rho(T_2) \le \kappa$. Continuity of the spectral radius extends this bound to $u = 0$ and $u = 1$. The displayed vector need not remain positive at those endpoints.

At $u = 0$, $D = 2$ and $E = 1$. The first two coordinates of $T_2(D, E, 0)$ give the block

$$\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix},$$

with eigenvalues $1 \pm \sqrt{2}$. The other four eigenvalues are zero. Thus the spectral radius is κ. This endpoint is attained by $(a, b, c) = (1, 1, 0)$, for which $L = 1$. Finally, restoring the original weights multiplies every transfer entry by L^2. □

5.3 A weighted trace bound

The next estimate uses two comparison vectors to control a trace. It gives a smaller finite range for the coefficient proof. For a positive vector ρ, define

$$F_\rho(B) = \sum_{i,j} \frac{\rho_i}{\rho_j} B_{ij}^2.$$

Lemma 5.4. *Let M be a finite square matrix with nonnegative entries. Suppose $K \geq 0$ and u, v have strictly positive coordinates, with $Mv \leq Kv$ and $M^{\mathsf{T}}u \leq Ku$ coordinatewise. Put $\rho_i = u_i/v_i$. For every integer $n \geq 4$,*

$$\operatorname{tr}(M^n) \leq K^{n-4} F_\rho(M^2).$$

Proof. Write $\|x\|_\rho^2 = \sum_i \rho_i x_i^2$. Weighted Cauchy–Schwarz gives

$$\left(\sum_j M_{ij} x_j\right)^2 \leq \left(\sum_j M_{ij} v_j\right)\left(\sum_j M_{ij} \frac{x_j^2}{v_j}\right).$$

Multiply by u_i/v_i, sum over i, and use the two comparison bounds. Then

$$\|Mx\|_\rho^2 \leq K \sum_j \frac{x_j^2}{v_j} \sum_i u_i M_{ij} \leq K^2 \|x\|_\rho^2.$$

Applying this estimate to each column gives $F_\rho(MB) \leq K^2 F_\rho(B)$. Thus, for $n \geq 4$,

$$F_\rho(M^{n-2}) \leq K^{2(n-4)} F_\rho(M^2).$$

A second Cauchy–Schwarz inequality, applied to the paired entries X_{ij} and Y_{ji}, gives

$$|\operatorname{tr}(XY)|^2 \leq F_\rho(X) F_\rho(Y).$$

Take $X = M^{n-2}$ and $Y = M^2$. All quantities needed to take the square root are nonnegative. This proves the result, including $K = 0$. □

Apply the lemma to the normalized transfer matrices. The bounds in (27) give $0 \leq f \leq 1$, $d \leq D(f) = 2 - 2f + 2f^2$, and $e \leq E(f) = 1 - f^2$. Set

$$K = \frac{169}{70}.$$

For the four-state matrix, take $w = (8, 4f, 2f^2, f^3)^{\mathsf{T}}$ when $f > 0$. The last three coordinates of $T_1 w$ equal those of $2w$. At $d = D(f)$ and $e = E(f)$, the first residual is

$$2w_1 - (T_1 w)_1 = 8(1 - f) + 6f^2 + 8f^2(1 - f) + f^4 \geq 0.$$

Smaller nonnegative d, e preserve this inequality. At $f = 0$, take the positive vector $w = (1, 1/4, 1/4, 1/4)^{\mathsf{T}}$. It satisfies $T_1 w \leq 2w$. Conjugation by the diagonal matrix with diagonal w gives a nonnegative matrix with row sums at most two. Its nth power has row sums at most 2^n. Each diagonal entry is at most this row sum, so

$$\operatorname{tr}(T_1^n) \leq 4 \cdot 2^n. \tag{31}$$

For $0 < f < 1$, put $D = D(f)$ and $E = E(f)$ and take

$$v = \left(1, \frac{E}{K - f}, \frac{f}{K}, \frac{E^2}{K(K - f)} + \frac{Df}{K^2} + \frac{f^3}{K^3}, \frac{fE}{K(K - f)}, \frac{f^2}{K^2}\right)^{\mathsf{T}}.$$

All six coordinates are positive. Rows two through six satisfy $(T_2 v)_i = K v_i$ at (D, E, f). The first residual is $R(f)/(K^3(K - f))$, where

$$R(f) = \frac{1}{1680700000} \sum_{j=0}^{5} \binom{5}{j} \gamma_j f^j (1 - f)^{5-j}$$

and

$$(\gamma_0, \ldots, \gamma_5) = (4826809, 18926317943, 24468294383, 26091136149, 22171767241, 3312639).$$

Every coefficient is positive, so $T_2 v \le K v$. Entrywise comparison gives the same result for $d \le D$ and $e \le E$. Let u be v with coordinates three and four exchanged. The literal matrix (24) satisfies $T_2^{\mathsf{T}} u \le K u$ by the same coordinate exchange.

It remains to bound $F_\rho(T_2^2)$, where $\rho_i = u_i / v_i$. Put $r = v_3/v_4$, so $\rho = (1, 1, r^{-1}, r, 1, 1)$. Direct multiplication gives

$$\begin{aligned} F_\rho(T_2^2) &= E_0(d,e,f) + E_1(d,e,f) r + (e^2+f^2)^2 r^2 \\ &\quad + \frac{2f^2(d^2+e^2)}{r} + \frac{f^4}{r^2}, \\ E_0(d,e,f) &= d^4 + 4d^2e^2 + 2e^4 + 4de^2 f + 6e^2 f^2 + 4f^4, \\ E_1(d,e,f) &= 2d^2 f^2 + 4de^2 f + 2e^4 + 2e^2 f^2. \end{aligned} \tag{32}$$

For fixed positive r and nonnegative f, this expression increases with $d, e \ge 0$. It is enough to use D, E. Define

$$\begin{aligned} N(f) &= 28561 f(169 - 70f), \\ H(f) &= 70\big[28561 E(f)^2 + 169 D(f) f(169 - 70f) + 70 f^3 (169 - 70f)\big]. \end{aligned}$$

The displayed vector gives $r = N(f)/H(f)$. Both $N(f)$ and $H(f)$ are positive for $0 < f < 1$. Let $G(d, e, f, r)$ denote the right side of (32). The exact polynomial identity

$$\big[39 - G(D(f), E(f), f, N(f)/H(f))\big] N(f)^2 H(f)^2 = f^2 \sum_{j=0}^{18} \beta_j f^j (1-f)^{18-j} \tag{33}$$

has the following coefficients.

j	β_j
0	46562161248457958625062 4500
1	17217514319785244183937758540
2	22365773062316640943215862 4159
3	15222959280418413298666559 41164
4	64421017408978916029979506 85822
5	18750207471127803372365770 643628
6	39927856862499630218557168 321919
7	64748235102644890096432222 301224
8	82112200779796853177563019 333638
9	82873919498476120791804987 635104
10	67279400669547001480535238 777620
11	44117805232440322189064060 340720
12	23297875076186045893202605 884378
13	97923134554622604994181079 93748
14	31960804797780734385941864 90087
15	77284904823008845233900890 4612
16	12590906976571730758907265 6246
17	10768695141523513153718037 708
18	63915177105473811509108079

All coefficients are positive. Thus (33) proves $F_\rho(T_2^2) \le 39$ in the open interval.

At the endpoints, use the positive vectors

$$v^{(0)} = (1, K-2, 1/9800, 1, 1/9800, 1/9800)^{\mathsf{T}},$$
$$v^{(1)} = (1, 1, 1/K, 2/K^2 + 1/K^3, 1/K, 1/K^2)^{\mathsf{T}}.$$

Direct multiplication verifies $T_2 v^{(i)} \le K v^{(i)}$ at $(D(i), E(i), i)$. The reflected left vectors satisfy the corresponding transpose bounds. At zero, $r = 1/9800$, and (32) is $34 + 2/9800 + 1/9800^2 < 39$. At one, $r = 28561/28560$, and the expression is $20 + 8(r + r^{-1}) + r^2 + r^{-2} < 39$. For the latter inequality, use $1 \le r \le 101/100$ and $r^{-1} \le 1$. This proves the energy bound at both endpoints without continuity. Lemma 5.4 now gives

$$\operatorname{tr}(T_2^n) \le 39K^{n-4}, \qquad n \ge 4. \tag{34}$$

5.4 The analytic bound from order 23

Combining (25), (31), and (34), then restoring the scale L, gives

$$\operatorname{per}(C_n C_n^{\mathsf{T}}) \le B_n L^{2n}, \qquad B_n = 2 + 8 \cdot 2^n + 39K^{n-4}, \qquad n \ge 4. \tag{35}$$

The exact integer inequality

$$23!\left[(2 + 8 \cdot 2^{23})70^{19} + 39 \cdot 169^{19}\right] < 23^{23}70^{19}$$

proves $c_{23}B_{23} < 1$. Since $K > 2$, $B_{n+1} \le KB_n$ for $n \ge 4$. The binomial theorem gives, for $n \ge 23$,

$$\left(1 + \frac{1}{n}\right)^n \ge 2 + \frac{n-1}{2n} \ge \frac{57}{23} > \frac{169}{70} = K.$$

Also, $c_{n+1}/c_n = (n/(n+1))^n$. Therefore $c_{n+1}B_{n+1} < c_n B_n$, and induction yields

$$c_n B_n < 1 \qquad (n \ge 23). \tag{36}$$

If at least one weight is positive, then $L > 0$. Equations (19), (35), and (36) give $c_n \operatorname{per}(C_n C_n^{\mathsf{T}}) < \operatorname{per}(C_n)^2$. If all weights are zero, the weak inequality holds directly.

5.5 An additional analytic bound from order 48

For any square matrix M of dimension m,

$$|\operatorname{tr}(M^n)| \le m\rho(M)^n.$$

This follows by summing powers of the eigenvalues with their algebraic multiplicities. It does not require diagonalizability. Since $f = ac \le L^2$, Lemma 5.3 and (25) give

$$\operatorname{per}(C_n C_n^{\mathsf{T}}) \le (2 + 14\kappa^n)L^{2n} \le 16\kappa^n L^{2n}. \tag{37}$$

For nonzero weights, Lemma 5.2 therefore gives

$$\frac{c_n \operatorname{per}(C_n C_n^{\mathsf{T}})}{\operatorname{per}(C_n)^2} \le 16 c_n \kappa^n. \tag{38}$$

The bound is strictly less than one at $n = 48$. This uses the exact integer comparison

$$16 \cdot 48! \cdot 169^{48} < 3360^{48}. \tag{39}$$

Indeed, $\sqrt{2} < 99/70$ because $99^2 > 2 \cdot 70^2$. Thus $\kappa < 169/70$, and $3360 = 48 \cdot 70$. Equation (39) implies $16c_{48}\kappa^{48} < 1$ without a floating-point estimate.

To extend this bound to all larger orders, put $t_n = 16c_n\kappa^n$. Then

$$\frac{t_{n+1}}{t_n} = \frac{\kappa}{(1+1/n)^n}.$$

The denominator increases with n. For completeness, the derivative of $x\log(1+1/x)$ is $\log(1+1/x) - 1/(x+1) > 0$ for $x > 0$. The last inequality follows from $\log(1+t) > t/(1+t)$ for $t > 0$. Moreover, $\kappa < 5/2$ and the integer comparison $5 \cdot 6^6 < 2 \cdot 7^6$ gives $(7/6)^6 > 5/2$. Consequently, $t_{n+1} < t_n$ for every $n \geq 6$. The first Gram bound is therefore strict for a nonzero cyclic matrix in every order $n \geq 48$.

5.6 Exact polynomial certificates for orders 3 through 22

The remaining finite range uses the homogeneous integer polynomial

$$\mathcal{D}_n(a,b,c) = n^n \operatorname{per}(C_n)^2 - n! \operatorname{per}(C_n C_n^{\mathsf{T}}). \tag{40}$$

Its degree is $2n$. For each ordering of the weights, write the ordered values as

$$x, \qquad x+y, \qquad x+y+z, \qquad x,y,z \geq 0.$$

These are gap coordinates. Every nonnegative triple belongs to at least one such closed region. Reflection of the cyclic indices exchanges a and c and preserves both permanents in (40). It is therefore enough to check the three substitutions

$$\begin{aligned} (a,b,c) &= (x, x+y, x+y+z), \\ (a,b,c) &= (x+y, x, x+y+z), \\ (a,b,c) &= (x, x+y+z, x+y). \end{aligned} \tag{41}$$

Their reflections give the other three orderings. Equal weights and zero weights are included in the boundaries.

The certificate computes every coefficient of each substituted $\mathcal{D}_n$. The following recurrences give the explicit connection to the matrix inequality. Direct determinant expansion of (23) and (24) gives

$$\begin{aligned} \det(tI - T_1) &= t^4 - et^3 - dft^2 - ef^2t - f^4, \\ \det(tI - T_2) &= t^6 - dt^5 - (e^2+f^2)t^4 - 2e^2ft^3 \\ &\qquad - (e^2f^2 + f^4)t^2 + df^4t + f^6. \end{aligned} \tag{42}$$

For a characteristic polynomial $t^m + h_1 t^{m-1} + \cdots + h_m$, let s_n be the trace of the nth power. Newton's identities give $s_0 = m$ and

$$\begin{aligned} s_n + \sum_{j=1}^{n-1} h_j s_{n-j} + nh_n &= 0 \quad (1 \leq n \leq m), \\ s_n + \sum_{j=1}^{m} h_j s_{n-j} &= 0 \quad (n > m). \end{aligned} \tag{43}$$

These identities compute $\operatorname{tr}(T_1^n)$ and $\operatorname{tr}(T_2^n)$ with integer polynomial arithmetic. Equation (25) then computes the Gram permanent.

The square of the cyclic permanent also has integer recurrences. Use $f = ac$ and define

$$U_n = \lambda^{2n} + \nu^{2n}, \qquad V_n = (a\lambda)^n + (a\nu)^n, \qquad W_n = (c\lambda)^n + (c\nu)^n.$$

Their initial values and recurrences are

$$\begin{aligned} U_0 &= 2, \quad U_1 = b^2 + 2f, \quad U_n = (b^2 + 2f)U_{n-1} - f^2 U_{n-2}, \\ V_0 &= 2, \quad V_1 = ab, \qquad\quad V_n = abV_{n-1} + a^2 f V_{n-2}, \\ W_0 &= 2, \quad W_1 = cb, \qquad\quad W_n = cbW_{n-1} + c^2 f W_{n-2} \qquad (n \geq 2). \end{aligned} \tag{44}$$

These follow from $\lambda + \nu = b$ and $\lambda\nu = -f$. Expansion of (19) yields

$$\operatorname{per}(C_n)^2 = a^{2n} + c^{2n} + U_n + 2V_n + 2W_n + \begin{cases} 4f^n, & n \text{ even}, \\ 0, & n \text{ odd}. \end{cases} \tag{45}$$

The parity term is $2f^n + 2(-f)^n$. Thus the calculation does not use a numerical square root.

The program `independent_finite.mjs` implements these recurrences with native JavaScript `BigInt` integers. It constructs the weight polynomials directly in the three gap regions (41). For a homogeneous polynomial of degree d, it stores the coefficient of $x^{d-i-j}y^i z^j$ at each pair $i, j \geq 0$ with $i + j \leq d$. The array therefore contains every one of the $\binom{d+2}{2}$ coefficient positions, including zeros. Addition is entrywise, and multiplication is exact coefficient convolution. The program applies (43) and (44), then forms (40). It checks the sign of every coefficient and stops with an error if any coefficient is negative.

The formal proof uses orders 3 through 22. The three substitutions give $3 \cdot 20 = 60$ closed regions and

$$3 \sum_{n=3}^{22} \binom{2n+2}{2} = 25\,050$$

possible coefficient positions. Lean proves the exact polynomial identities and the nonnegativity of their coefficients. A separate integer computation extends through order 47 and gives

$$3 \cdot 45 = 135 \text{ regions}, \qquad 3 \sum_{n=3}^{47} \binom{2n+2}{2} = 224\,550 \text{ coefficient positions}.$$

Every coefficient is nonnegative. For $4 \leq n \leq 47$, every coefficient is positive. Nonnegativity of these coefficients proves $\mathcal{D}_n \geq 0$ throughout each gap region. Indeed, each monomial is nonnegative when $x, y, z \geq 0$. The three regions and their reflections cover all nonnegative real triples. This polynomial certificate covers all nonnegative real weights in the stated finite range.

The program and its recorded output are included in `papers/mathematics/gyires/research-notes/extensions/` in the accompanying repository. The program is in `certificates/independent_finite.mjs`. The record is in `results/independent_finite.json`. From a new working directory below `gyires-extensions/`, the certificate command is

```
node ../certificates/independent_finite.mjs 47
```

It writes the coefficient-count record to the working directory. The associated `verification.md` file gives the record checks and hashes.

Two further programs check the supporting identities. The program `sharp_bound_check.py` verifies the residual factorization in $\mathbb{Q}(\sqrt{2})$ and the integer comparison (39). The program `circ_small_check.py` computes permanents by subsets. Its recorded outputs agree with both transfer formulas for $1 \leq n \leq 7$. It also checks all six gap orderings in orders three and four. These checks support the transfer implementation. The proof for unbounded orders uses the analytic bound and its monotone continuation established above.

5.7 Small orders, boundary cases, and permutation cycles

For $n = 1$, C_1 is the scalar $a + b + c$, and the Gram inequality is equality. For $n = 2$, put $x = b$ and $y = a + c$. Then

$$C_2 = \begin{pmatrix} x & y \\ y & x \end{pmatrix}, \qquad \operatorname{per}(C_2) = x^2 + y^2, \qquad \operatorname{per}(C_2 C_2^{\mathsf{T}}) = (x^2 + y^2)^2 + 4x^2y^2.$$

Since $c_2 = 1/2$,

$$\operatorname{per}(C_2)^2 - \tfrac{1}{2}\operatorname{per}(C_2 C_2^{\mathsf{T}}) = \tfrac{1}{2}(x^2 - y^2)^2 \geq 0.$$

For orders three through 22, the closed gap regions include all zero-weight boundaries. For every order at least 23, the weighted trace argument includes both normalized endpoints through explicit positive vectors. If $a = b = c = 0$, both permanents vanish in every order. Thus every cyclic case is covered.

Proof of Theorem 5.1. Suppose the cycles of Q have lengths $m_1, \ldots, m_s$, where $\sum_i m_i = n$. A simultaneous row and column permutation writes the matrix C in (17) as a direct sum of cyclic matrices C_i. Permanents multiply over direct sums, as do the permanents of their Gram matrices. The cyclic inequalities give

$$\operatorname{per}(C)^2 \geq \left(\prod_{i=1}^{s} c_{m_i}\right) \operatorname{per}(CC^{\mathsf{T}}). \tag{46}$$

The multinomial theorem gives

$$1 = \left(\sum_{i=1}^{s} \frac{m_i}{n}\right)^n \geq \frac{n!}{\prod_i m_i!} \prod_{i=1}^{s} \left(\frac{m_i}{n}\right)^{m_i}.$$

Rearrangement yields $\prod_i c_{m_i} \geq c_n$. Together with (46), this proves $c_n \operatorname{per}(CC^{\mathsf{T}}) \leq \operatorname{per}(C)^2$. Equation (17) proves the first bound for A.

The transpose $A^{\mathsf{T}} = aI + bQ^{\mathsf{T}} + c(Q^{\mathsf{T}})^2$ has the same form. The first bound applied to A^{T} proves the second bound in (15). The permanent Gram comparison, applied to A and A^{T}, gives

$$2\operatorname{per}(A^2) \leq \operatorname{per}(AA^{\mathsf{T}}) + \operatorname{per}(A^{\mathsf{T}}A).$$

The two Gram bounds therefore prove the third bound in the original indexing of A. Adding the three bounds with weights $1, 1, 2$ proves (16) for all nonnegative weights. If $a + b + c = 1$, each row sum and each column sum of A is one. □

Corollary 5.5. *Suppose $Q = P_n$, $n \geq 23$, and at least one of a, b, c is positive. Then all three inequalities in* (15) *and the combined inequality* (16) *are strict.*

Proof. The strict weighted trace estimate proves the first Gram bound. Since A is a polynomial in Q and $Q^{\mathsf{T}} = Q^{-1}$, the two Gram matrices are equal. Thus the second Gram bound is also strict. The permanent Gram comparison and addition give the strict square and combined bounds. □

The theorem concerns a quadratic polynomial in one permutation matrix. It does not assert the same result for arbitrary triples of permutation matrices or arbitrary circulants.

6 Four-block matrices in arbitrary dimensions

The next result concerns matrices with at most two row types and at most two column types. Two rows have the same type if their entries agree in every column. Column types have the corresponding meaning. The block sizes can be unequal, and the entries can be zero.

Theorem 6.1 (Four-block theorem)**.** *Let $n \geq 2$. Let A be a nonnegative doubly stochastic $n \times n$ matrix with at most two distinct rows and at most two distinct columns. Set $d = \operatorname{per}(A)$. Then*

$$\operatorname{per}(AA^{\mathsf{T}}) \leq d, \qquad \operatorname{per}(A^{\mathsf{T}}A) \leq d, \qquad \operatorname{per}(A^2) \leq d. \tag{47}$$

Moreover,

$$\begin{aligned} 4d^2 - c_n\left[\operatorname{per}(AA^{\mathsf{T}}) + \operatorname{per}(A^{\mathsf{T}}A) + 2\operatorname{per}(A^2)\right] &\geq 4d(d - c_n) \\ &\geq 0. \end{aligned} \tag{48}$$

Equality in Gyires's inequality holds if and only if $A = J_n$.

Here A^2 denotes the matrix product. The expression d^2 denotes the square of the number $\operatorname{per}(A)$. The equality statement concerns Gyires's inequality. It does not assert that equality in any one of (47) forces $A = J_n$.

6.1 The four blocks and their full parameter interval

If A has one row type, each column sum forces every entry in that column to be $1/n$. Thus $A = J_n$. One column type gives the same conclusion from the row sums.

Suppose that both types occur. Let the row classes have sizes r and $n - r$. Let the column classes have sizes s and $n - s$, where $1 \leq r, s \leq n - 1$. Every entry in a given block is equal. Let h be the total mass in the block with r rows and s columns. The four block masses are

$$h, \qquad r - h, \qquad s - h, \qquad n - r - s + h.$$

After the row classes and the column classes are placed together, the four block entries are therefore

$$\begin{pmatrix} \dfrac{h}{rs} & \dfrac{r-h}{r(n-s)} \\[2ex] \dfrac{s-h}{(n-r)s} & \dfrac{n-r-s+h}{(n-r)(n-s)} \end{pmatrix}. \tag{49}$$

Each displayed entry fills the corresponding block. Nonnegativity is equivalent to the closed interval

$$\max(0, r + s - n) \leq h \leq \min(r, s). \tag{50}$$

Conversely, each h in this interval gives a doubly stochastic matrix. For example, a row in the first class has sum $h/r + (r - h)/r = 1$. The other row sums and both column sums follow in the same way. The uniform matrix occurs exactly when $h = rs/n$.

The grouping in (49) describes the matrix. It does not replace the original indexing when A^2 is formed. Independent row and column permutations need not preserve $\operatorname{per}(A^2)$. The proof below treats this term through Lemma 2.2, in the original indexing.

The main estimate first uses only the condition on column types. It follows from a probability inequality, proved next.

6.2 A probability inequality at an integer mean

A Bernoulli variable with parameter $p \in [0,1]$ has value 1 with probability p and value 0 with probability $1-p$. A sum of independent Bernoulli variables can have different parameters. Its distribution is often called a Poisson-binomial distribution.

The proof starts with the required fixed-mean maximizer principle. This is the classical reduction of Hoeffding [12]. The proof includes all boundary parameters.

Lemma 6.2 (Fixed-mean maximizer)**.** *Let $\mu \in [0,n]$, and let $f : \{0,\dots,n\} \to \mathbb{R}$ be any function. Among all independent Bernoulli sums K with parameters $p_1,\dots,p_n \in [0,1]$ and $\sum_i p_i = \mu$, the maximum of $\mathbb{E}f(K)$ occurs at parameters whose values in $(0,1)$ are all equal. Thus a maximizing distribution is either deterministic or has the form*

$$K = a + \mathrm{Bin}(b,q), \qquad a \ge 0, \quad b \ge 1, \quad a+b \le n, \quad 0 < q < 1, \quad q = \frac{\mu - a}{b},$$

where a and b are integers.

Proof. The parameter set

$$\mathcal{P}_\mu = \{(p_1,\dots,p_n) \in [0,1]^n : \textstyle\sum_i p_i = \mu\}$$

is nonempty and compact. The objective $\mathbb{E}f(K)$ is a polynomial in the parameters, so it attains a maximum. Among its maximizers, choose one that also maximizes $\sum_i p_i^2$. The set of maximizers is compact, so this second maximum exists.

Suppose that two distinct parameters p, q lie in $(0,1)$. Keep their sum $t = p+q$ fixed, and write $z = pq$. Their contribution to the probability generating polynomial is

$$((1-p)+px)((1-q)+qx) = (1-t+z) + (t-2z)x + zx^2.$$

The other factors are fixed. Hence $\mathbb{E}f(K)$ is an affine function of z. The feasible interval for z is

$$\max(0, t-1) \le z \le t^2/4.$$

The upper endpoint corresponds to $p = q$. The lower endpoint has one parameter equal to 0 or 1. Since p and q are distinct and interior, the present z is strictly between the endpoints. An affine function with a maximum at an interior point is constant. Moving z to the lower endpoint therefore does not change the objective. This move strictly increases $p^2+q^2 = t^2-2z$. That contradicts the second maximization.

All interior parameters at the chosen maximizer are consequently equal. Let a parameters equal 1, let b parameters lie in $(0,1)$, and let the remaining parameters equal 0. If $b = 0$, the sum is deterministic. If $b > 0$, its distribution is $a + \mathrm{Bin}(b,q)$, and its mean gives $a + bq = \mu$. □

Lemma 6.3 (Adjacent probabilities)**.** *Let $P_k = \mathbb{P}(K = k)$ for a sum of n independent Bernoulli variables. Suppose that its mean is an integer m with $1 \le m \le n-1$. Then*

$$(m+1)P_{m+1} \le mP_m, \qquad (n-m+1)P_{m-1} \le (n-m)P_m. \tag{51}$$

Proof. First maximize the linear functional $(m+1)P_{m+1} - mP_m$ over the parameters with sum m. Lemma 6.2 gives a maximizing distribution of the stated form. A deterministic distribution with mean m has $P_m = 1$ and $P_{m+1} = 0$, so this functional is negative.

For a nondeterministic maximizer, write $K = a + \mathrm{Bin}(b,q)$ and $u = m - a$. The mean condition gives $q = u/b$. Since m and a are integers,

$$1 \le u \le b-1, \qquad u \le m.$$

In particular, $P_m > 0$. The adjacent binomial ratio gives

$$\frac{P_{m+1}}{P_m} = \frac{b-u}{u+1}\frac{q}{1-q} = \frac{u}{u+1} \le \frac{m}{m+1}.$$

The maximum of the functional is therefore nonpositive. This proves the first inequality for every parameter vector.

Next maximize $(n-m+1)P_{m-1} - (n-m)P_m$ in the same parameter set. The deterministic case again gives a negative value. For a nondeterministic maximizer, use the same notation. Since $a+b \le n$, $b-u = b-m+a \le n-m$. It follows that

$$\frac{P_{m-1}}{P_m} = \frac{u}{b-u+1}\frac{1-q}{q} = \frac{b-u}{b-u+1} \le \frac{n-m}{n-m+1}.$$

This proves the second inequality. □

Proposition 6.4 (Central likelihood-ratio bound)**.** *Under the assumptions of Lemma 6.3, put*

$$\theta = \frac{m}{n}, \qquad B_k = \binom{n}{k}\theta^k(1-\theta)^{n-k}, \qquad R_k = \frac{P_k}{B_k} \quad (0 \le k \le n).$$

Then $B_k > 0$, $P_m > 0$, and

$$R_k \le R_m \quad (0 \le k \le n). \tag{52}$$

Consequently,

$$\sum_{k=0}^{n} \frac{P_k^2}{B_k} \le \frac{P_m}{B_m}, \qquad \frac{P_m}{B_m} \ge 1. \tag{53}$$

Equality $P_m/B_m = 1$ holds if and only if every Bernoulli parameter equals m/n.

Proof. Since $0 < \theta < 1$, every B_k is positive. The probability generating polynomial is

$$G(x) = \prod_{i=1}^{n}(1-p_i+p_ix) = \sum_{k=0}^{n} P_k x^k.$$

When all p_i lie in $(0,1)$, this polynomial has degree n and all its roots are real and negative. The classical Newton inequalities give

$$\left(\frac{P_k}{\binom{n}{k}}\right)^2 \ge \frac{P_{k-1}}{\binom{n}{k-1}}\frac{P_{k+1}}{\binom{n}{k+1}} \qquad (1 \le k \le n-1). \tag{54}$$

Each coefficient depends continuously on the parameters. Approximation by vectors in $(0,1)^n$ therefore proves (54) also when parameters equal 0 or 1. The approximating vectors need not have mean m, since Newton's inequalities hold without a mean constraint.

An inequality of the form $v_k^2 \ge v_{k-1}v_{k+1}$ is called log-concavity. Multiplication by a positive geometric sequence preserves this property. Thus (54) implies that $(R_k)_{k=0}^n$ is log-concave, because

$$R_k = \frac{P_k}{\binom{n}{k}}(1-\theta)^{-n}\left(\frac{1-\theta}{\theta}\right)^k.$$

To locate its positive terms, let a parameters equal 1 and let b parameters lie in $(0,1)$. The remaining parameters equal 0. If $b = 0$, then $K = m$ deterministically. If $b > 0$, every integer in $\{a, \dots, a+b\}$ has positive probability. No other integer does. Also,

$$a < m = a + \sum_{0<p_i<1} p_i < a+b.$$

In both cases $P_m > 0$, and the support is an interval of integers.

The binomial probabilities at the mean satisfy

$$\frac{B_{m+1}}{B_m} = \frac{m}{m+1}, \qquad \frac{B_{m-1}}{B_m} = \frac{n-m}{n-m+1}.$$

Lemma 6.3 therefore gives $R_{m-1} \le R_m$ and $R_{m+1} \le R_m$. On the positive support of a log-concave sequence, the successive ratios R_{k+1}/R_k are nonincreasing. The first adjacent inequality implies that all ratios to its left are at least 1. The second implies that all ratios to its right are at most 1. Hence R_m is a maximum on the positive support. Outside the support, $R_k = 0$. The deterministic case has only the single positive term R_m. This proves (52) in all cases.

The first inequality in (53) now follows from

$$\sum_{k=0}^{n} \frac{P_k^2}{B_k} = \sum_{k=0}^{n} P_k R_k \le R_m \sum_{k=0}^{n} P_k = R_m.$$

The second follows from

$$1 = \sum_{k=0}^{n} B_k R_k \le R_m \sum_{k=0}^{n} B_k = R_m.$$

If $R_m = 1$, then $R_k \le 1$ for all k. Since every $B_k > 0$ and their weighted average is 1, every R_k must equal 1. Thus $P_k = B_k$ for all k, and

$$\prod_{i=1}^{n} (1 - p_i + p_i x) = (1 - \theta + \theta x)^n. \tag{55}$$

The right side has nonzero constant term and degree n. The left side therefore has no parameter equal to 1 or 0. Its linear factors have roots $-(1-p_i)/p_i$. The only root on the right side is $-(1-\theta)/\theta$. Uniqueness of polynomial factorization gives $p_i = \theta$ for every i. Conversely, these parameters give $P = B$ and $R_m = 1$. □

Corollary 6.5 (Equality in the stronger probability bound)**.** *Under the assumptions of Proposition 6.4, equality in*

$$\sum_k \frac{P_k^2}{B_k} \le \frac{P_m}{B_m}$$

holds exactly when $P = B$ or P is the point mass at m. Equality in the weaker bound with right side $(P_m/B_m)^2$ holds exactly when $P = B$.

Proof. The first bound is an average of $R_k \le R_m$ with weights P_k. Equality therefore forces $R_k = R_m$ at every point of the positive support. Thus P is the restriction of B to an interval $[a, b]$, followed by normalization. Put $d = b - a$. If $d \ge 2$, remove the factor z^a from the probability generating polynomial. The resulting degree-d polynomial still has only real nonpositive roots. At $1 \le j \le d-1$, the ratio of three adjacent coefficients inherited from B is

$$\frac{P_{a+j}^2}{P_{a+j-1} P_{a+j+1}} = \frac{a+j+1}{a+j} \frac{n-a-j+1}{n-a-j}.$$

Newton's inequality for the actual degree d requires this ratio to be at least

$$\frac{j+1}{j} \frac{d-j+1}{d-j}.$$

Since $a+j \geq j$ and $n-a-j \geq d-j$, each factor in the first display is at most its counterpart in the second. At least one is strictly smaller unless $a=0$ and $d=n$. A proper interval of width at least two is therefore impossible. An interval of width one with positive mass at both endpoints has noninteger mean. Width zero gives the point mass at m; the full interval gives $P=B$. Both attain the first equality. For the point mass, $P_m/B_m = 1/B_m > 1$. Proposition 6.4 now gives the equality statement for the squared bound. □

6.3 From two column types to a permanent contraction

Proposition 6.6 (Two-column contraction)**.** *Let A be a nonnegative doubly stochastic $n \times n$ matrix with at most two column types. Then*

$$\operatorname{per}(AA^{\mathsf{T}}) \leq \operatorname{per}(A), \qquad \operatorname{per}(A) \geq c_n. \tag{56}$$

Equality $\operatorname{per}(A) = c_n$ *holds if and only if* $A = J_n$.

Proof. One column type forces $A = J_n$, for which the assertions hold. Otherwise let I be a column class of size s, where $1 \leq s \leq n-1$. Let u_i be the total mass in row i over the columns in I. Because columns within a class agree, row i has entry u_i/s on I and entry $(1-u_i)/(n-s)$ off I. Nonnegativity and the row sums give $0 \leq u_i \leq 1$. The column sums give

$$\sum_{i=1}^{n} u_i = s.$$

Let P_k be the probability that a sum of independent Bernoulli variables with parameters $u_1, \ldots, u_n$ equals k. Let B_k be the probabilities of $\operatorname{Bin}(n, s/n)$.

For variables $z_1, \ldots, z_n$, set

$$X = \sum_{j \in I} z_j, \qquad Y = \sum_{j \notin I} z_j, \qquad p_A(z) = \prod_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} z_j \right).$$

The Bernoulli generating function gives the exact expansion

$$p_A(z) = \prod_{i=1}^{n} \left(\frac{u_i}{s} X + \frac{1-u_i}{n-s} Y \right) = \sum_{k=0}^{n} \frac{P_k}{s^k (n-s)^{n-k}} X^k Y^{n-k}. \tag{57}$$

The coefficient of $z_1 \cdots z_n$ is $\operatorname{per}(A)$. In (57), only the term $k=s$ can contribute to this coefficient. The coefficient of $\prod_{j \in I} z_j$ in X^s is $s!$. The coefficient from Y^{n-s} is $(n-s)!$. Consequently,

$$\operatorname{per}(A) = \frac{s!(n-s)!}{s^s (n-s)^{n-s}} P_s = c_n \frac{P_s}{B_s}. \tag{58}$$

The last equality uses

$$B_s = \frac{n!}{s!(n-s)!} \frac{s^s (n-s)^{n-s}}{n^n}.$$

The factorial coefficient inner product follows from Lemma 2.1. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_s)$, write $|\alpha| = \sum_j \alpha_j$ and $\alpha! = \prod_j \alpha_j!$. The coefficient of a monomial with multi-index α in X^k is $k!/\alpha!$ when $|\alpha| = k$. Thus its squared Fischer norm is

$$\sum_{|\alpha|=k} \alpha! \left(\frac{k!}{\alpha!} \right)^2 = (k!)^2 \sum_{|\alpha|=k} \frac{1}{\alpha!} = k! s^k. \tag{59}$$

The last equality is the multinomial theorem at s variables equal to 1. It includes $k = 0$. In the second column class the same calculation gives $(n-k)!(n-s)^{n-k}$. The two classes use disjoint variables, so these squared norms multiply. Terms with different k have different total degrees in the first class. They therefore have disjoint monomial supports and are orthogonal in the factorial inner product.

Lemma 2.1 and (57) now give

$$\begin{aligned}\operatorname{per}(AA^{\mathsf{T}}) &= \langle p_A, p_A\rangle_F \\ &= \sum_{k=0}^{n}\left(\frac{P_k}{s^k(n-s)^{n-k}}\right)^2 k!s^k(n-k)!(n-s)^{n-k} \\ &= \sum_{k=0}^{n}\frac{P_k^2 k!(n-k)!}{s^k(n-s)^{n-k}} = c_n\sum_{k=0}^{n}\frac{P_k^2}{B_k}. \end{aligned} \tag{60}$$

Here the last step uses, for each k,

$$B_k = \frac{n!}{k!(n-k)!}\frac{s^k(n-s)^{n-k}}{n^n}.$$

The Bernoulli sum has integer mean s. Proposition 6.4, (58), and (60) imply

$$\operatorname{per}(AA^{\mathsf{T}}) = c_n\sum_{k=0}^{n}\frac{P_k^2}{B_k} \le c_n\frac{P_s}{B_s} = \operatorname{per}(A), \qquad \operatorname{per}(A) \ge c_n.$$

If $\operatorname{per}(A) = c_n$, the equality statement in Proposition 6.4 gives $u_i = s/n$ for every i. Both entries in each row are then $1/n$, so $A = J_n$. The converse follows from $\operatorname{per}(J_n) = c_n$. □

6.4 Proof of the four-block theorem

Proof of Theorem 6.1. Proposition 6.6 gives $\operatorname{per}(AA^{\mathsf{T}}) \le d$ and $d \ge c_n$. The transpose A^{T} also has at most two column types, because A has at most two row types. Since $\operatorname{per}(A^{\mathsf{T}}) = \operatorname{per}(A)$, the same proposition gives $\operatorname{per}(A^{\mathsf{T}}A) \le d$. Lemma 2.2 gives

$$\operatorname{per}(A^2) \le \sqrt{\operatorname{per}(AA^{\mathsf{T}})\operatorname{per}(A^{\mathsf{T}}A)} \le d.$$

This proves (47) without any change to the indexing of A^2.

Adding these estimates with weights $1, 1, 2$ gives

$$\operatorname{per}(AA^{\mathsf{T}}) + \operatorname{per}(A^{\mathsf{T}}A) + 2\operatorname{per}(A^2) \le 4d.$$

Multiplication by $c_n > 0$ and subtraction from $4d^2$ give (48). If $A \ne J_n$, Proposition 6.6 gives $d > c_n$. Thus $4d(d - c_n) > 0$, and Gyires's inequality is strict. If $A = J_n$, then $A^2 = AA^{\mathsf{T}} = A^{\mathsf{T}}A = J_n$ and $d = c_n$. Both sides of Gyires's inequality are then $4c_n^2$. □

All endpoints in (50) are included. The probability proof allowed $u_i = 0$ and $u_i = 1$, so it requires no positive-entry assumption. The one-type case was included separately. For $n = 1$, the only doubly stochastic matrix is $J_1 = (1)$, and the same conclusions hold directly.

The individual contraction inequalities can have other equality cases. For $1 \le r \le n-1$, take $A = \operatorname{diag}(J_r, J_{n-r})$. This matrix has at most two row types and at most two column types, and $AA^{\mathsf{T}} = A^{\mathsf{T}}A = A^2 = A$. Thus all three contraction inequalities are equalities. However, $A \ne J_n$, so Gyires's inequality is strict. The theorem applies to the specified row and column types. It does not assert the result for every matrix of rank at most two.

7 Matrices with one exceptional column or row

7.1 Statement

Let A be a nonnegative doubly stochastic matrix of order $n \geq 2$ with $n-1$ identical columns. Then

$$\operatorname{per}(A)^2 \geq c_n \operatorname{per}(AA^T), \qquad \operatorname{per}(A)^2 \geq c_n \operatorname{per}(A^TA), \qquad c_n = \frac{n!}{n^n}. \tag{61}$$

Consequently Gyires's inequality holds for this family:

$$4\operatorname{per}(A)^2 \geq c_n \left(\operatorname{per}(AA^T) + \operatorname{per}(A^TA) + 2\operatorname{per}(A^2)\right). \tag{62}$$

Equality occurs exactly at $A = J_n$, where every entry of J_n is $1/n$. The same conclusions hold for matrices with $n-1$ identical rows.

The first Gram inequality follows from Proposition 6.6. The second has the elementary proof below.

7.2 Parametrization and the first Gram inequality

Move the special column to the first position. The permanent and both Gram permanents are unchanged by this column relabeling; the final Cauchy–Schwarz step can be applied to the original matrix as well. Every such matrix has columns

$$u,\ v, \ldots, v, \qquad u_i \geq 0, \quad \sum_i u_i = 1, \quad v_i = \frac{1-u_i}{m}, \quad m = n-1.$$

There are at most two column types. Proposition 6.6 gives

$$\operatorname{per}(AA^T) \leq \operatorname{per}(A), \qquad \operatorname{per}(A) \geq c_n.$$

Therefore $c_n \operatorname{per}(AA^T) \leq \operatorname{per}(A)^2$. If $A \neq J_n$, the strict permanent lower bound in that proposition makes this strong inequality strict. The second Gram inequality needs a separate argument, since A can have more than two row types.

7.3 An exact formula for the other Gram permanent

Put

$$a = \sum_i u_i^2, \qquad b = u \cdot v = \frac{1-a}{m}, \qquad c = \|v\|^2 = \frac{n-2+a}{m^2}.$$

Let E_m be the $m \times m$ all-ones matrix. Then

$$A^TA = \begin{pmatrix} a & b\mathbf{1}^T \\ b\mathbf{1} & cE_m \end{pmatrix}.$$

In its permanent, either the first row uses the first column, or it uses one of the other m columns. Counting the two cases gives

$$\operatorname{per}(A^T A) = m!\, c^{m-1}(ac + mb^2). \tag{63}$$

Let

$$Q(u) = \sum_i u_i \prod_{j \neq i} (1 - u_j)$$

be the probability of exactly one success among independent Bernoulli variables with parameters $u_1, \ldots, u_n$. The permanent is

$$\operatorname{per}(A) = \frac{m!}{m^m} Q(u). \tag{64}$$

After substituting (63) and (64), the second inequality in (61) is equivalent to

$$Q(u)^2 \geq R_n(a), \tag{65}$$

where

$$R_n(a) = n^{1-n}(n - 2 + a)^{n-2}\big(n - 1 - na + na^2\big), \qquad \frac{1}{n} \leq a \leq 1. \tag{66}$$

At $a_0 = 1/n$,

$$R_n(a_0) = g_n^2, \qquad g_n = \left(\frac{n-1}{n}\right)^{n-1}.$$

7.4 Two elementary lower bounds for the one-success probability

First, the product inequality

$$\prod_{j \neq i} (1 - u_j) \geq 1 - \sum_{j \neq i} u_j = u_i$$

gives

$$\boxed{Q(u) \geq a.} \tag{67}$$

For the second bound, initially suppose every $u_i < 1$. Weighted AM–GM, with weights u_i, gives

$$\log\left(\sum_i \frac{u_i}{1 - u_i}\right) \geq -\sum_i u_i \log(1 - u_i).$$

Using $Q = \prod_i (1 - u_i) \sum_i u_i/(1 - u_i)$,

$$\log Q \geq \sum_i (1 - u_i) \log(1 - u_i). \tag{68}$$

Let $f(x) = (1-x)\log(1-x)$, continuously extended by $f(1) = 0$. Since $f''(x) = 1/(1-x) \geq 1$, strong convexity at $1/n$ gives

$$\sum_i f(u_i) \geq nf(1/n) + \frac{1}{2}\sum_i (u_i - 1/n)^2 = \log g_n + \frac{1}{2}(a - 1/n).$$

Therefore

$$\boxed{Q(u)^2 \geq g_n^2 e^{a-1/n}.} \tag{69}$$

Both bounds extend to boundary points by continuity.

7.5 The range $1/n \leq a \leq 1/2$

Write

$$d = n - 2 + a, \qquad K = n - 1 - na + na^2.$$

Differentiating (66),

$$\frac{d}{da}\log R_n(a) = \frac{n-2}{d} + \frac{n(2a-1)}{K} \leq 1$$

throughout this range. Integrating from $1/n$ gives

$$R_n(a) \leq g_n^2 e^{a-1/n} \leq Q(u)^2. \tag{70}$$

For $n > 2$ and $a > 1/n$, the first inequality is strict. For $n = 2$, this range consists only of the uniform endpoint.

7.6 The range $1/2 \leq a \leq 1$

A direct algebraic simplification gives

$$\frac{d}{da}\log\frac{R_n(a)}{a^2} = \frac{(a-1)\left(n(n-2)a^2 + na + 2(n-1)(n-2)\right)}{a(n-2+a)(n-1-na+na^2)} \leq 0. \tag{71}$$

Thus it is enough to show $R_n(1/2) \leq 1/4$. At the endpoint,

$$R_n(1/2) = \left(1 - \frac{3}{2n}\right)^{n-2}\left(\frac{3}{4} - \frac{1}{n}\right).$$

For completeness, set $x = 3/(2n)$ and $k = n - 2$. The nonnegative binomial series gives

$$(1-x)^{-k} \geq 1 + kx + \frac{k(k+1)}{2}x^2 \geq 1 + \frac{2k}{n} = 3 - \frac{4}{n}.$$

The second inequality follows from

$$\frac{(k+1)x}{2} = \frac{3(n-1)}{4n} \geq \frac{1}{3}$$

for $n \geq 2$; the case $k = 0$ is equality directly. Hence $R_n(1/2) \leq 1/4$, with strict inequality for $n > 2$.

Combining this with (71) and (67),

$$R_n(a) \leq a^2 \leq Q(u)^2. \tag{72}$$

This proves the second Gram inequality in every order.

7.7 Equality and Gyires's inequality

The preceding bounds are strict unless $a = 1/n$. By $\sum_i u_i = 1$, that condition forces $u_i = 1/n$ for all i, so $A = J_n$.

Lemma 2.2, applied in the original indexing, gives

$$\operatorname{per}(A^2) \leq \sqrt{\operatorname{per}(AA^T)\operatorname{per}(A^TA)} \leq \frac{1}{2}\left(\operatorname{per}(AA^T) + \operatorname{per}(A^TA)\right).$$

Applying both inequalities in (61) now proves (62), with equality only at J_n. Transposition proves the corresponding statement when $n - 1$ rows are identical.

The identical-column hypothesis permits the exceptional column to have any number of distinct entries. It still excludes general rank-two matrices.

8 All rank-two matrices in orders two through fifteen

The four-block condition restricts the number of distinct rows and columns. A general matrix of rank at most two can have more row and column types. The next result covers all such doubly stochastic matrices in a finite range of orders. Its proof combines an analytic reduction with exact integer polynomial certificates.

Theorem 8.1 (Finite-order rank-two theorem)**.** *Let $2 \leq n \leq 15$, let $A \in \Omega_n$, and suppose* $\operatorname{rank}(A) \leq 2$. *Set $d = \operatorname{per}(A)$. Then*

$$\operatorname{per}(AA^{\mathsf{T}}) \leq d, \qquad \operatorname{per}(A^{\mathsf{T}}A) \leq d, \qquad \operatorname{per}(A^2) \leq d. \tag{73}$$

Moreover,

$$\Delta_n(A) \geq 4d(d - c_n) \geq 0, \qquad c_n \operatorname{per}(AA^{\mathsf{T}}) \leq d^2. \tag{74}$$

Equality in Gyires's inequality, or in the last strong inequality, holds exactly when $A = J_n$.

The contraction itself can have other equality cases. For example, a direct sum of two uniform stochastic blocks has rank two and satisfies $AA^{\mathsf{T}} = A$. The order-one case is immediate. Theorem 8.1 makes no claim for a general rank-two matrix of order greater than fifteen.

8.1 A representation of every rank-two matrix

A rank-one doubly stochastic matrix is J_n. Indeed, its rows are proportional, and their common row sum makes them equal. The column sums then determine every entry. Suppose that A is nonuniform and has rank at most two. Its row space has dimension at most two. The row-sum functional is nonzero on this space, so its intersection with the row-sum-one hyperplane has affine dimension at most one. The rows therefore lie on a line in the probability simplex.

Let $b^{(1)}$ and $b^{(0)}$ be the two endpoint rows of their convex hull. Write row i as $x_i b^{(1)} + (1-x_i)b^{(0)}$, where $0 \le x_i \le 1$. Both endpoints occur, so $t := \sum_i x_i$ satisfies $0 < t < n$. The column sums give

$$t b_j^{(1)} + (n-t) b_j^{(0)} = 1.$$

Set $y_j = t b_j^{(1)}$. Then $0 \le y_j \le 1$ and $\sum_j y_j = t$. This gives

$$A_{ij} = \frac{x_i y_j}{t} + \frac{(1-x_i)(1-y_j)}{n-t}, \qquad x, y \in [0,1]^n, \quad \sum_i x_i = \sum_j y_j = t. \tag{75}$$

Conversely, these conditions make (75) a nonnegative doubly stochastic matrix of rank at most two. Each summand has rank at most one. The row sum is $x_i + (1-x_i) = 1$, and the column sum follows in the same way.

For fixed n, t, define

$$\theta = \frac{t}{n}, \quad \sigma = \theta(1-\theta) > 0, \quad u = x - \theta \mathbf{1}, \quad v = y - \theta \mathbf{1}.$$

Then $\sum_i u_i = \sum_j v_j = 0$ and

$$A = J_n + \frac{u v^{\mathsf{T}}}{n\sigma}. \tag{76}$$

Let e_k denote the elementary symmetric polynomial of degree k, with $e_0 = 1$. Put

$$w_k = \frac{k!(n-k)!}{n! \sigma^k}, \qquad \kappa_y = \frac{\sum_j v_j^2}{n\sigma}.$$

Since $\sum_j y_j^2 \le \sum_j y_j = t$, $0 \le \kappa_y \le 1$.

Expanding the permanent by the rows and columns that use the rank-one term gives

$$S(x,y) := \frac{\operatorname{per}(A)}{c_n} = \sum_{k=0}^{n} w_k e_k(u) e_k(v). \tag{77}$$

For fixed sets of k rows and k columns, there are $k!$ matchings using that term. The other rows and columns contribute $(n-k)!/n^{n-k}$. The factor $(n\sigma)^{-k}$ and division by c_n give the displayed coefficient.

The zero sums of u, v make the cross terms with J_n vanish. Thus

$$A A^{\mathsf{T}} = J_n + \frac{\kappa_y}{n\sigma} u u^{\mathsf{T}},$$

and the same expansion yields

$$Q(x,y) := \frac{\operatorname{per}(A A^{\mathsf{T}})}{c_n} = \sum_{k=0}^{n} w_k e_k(u)^2 \kappa_y^k. \tag{78}$$

The zeroth term is one and the first term is zero in both expansions. Every term in (78) is nonnegative.

8.2 Separate convexity in the probability laws

Let P_x be the law of a sum of independent Bernoulli variables with parameters $x_1, \dots, x_n$. Its probability generating polynomial is

$$F_{P_x}(z) = \sum_{j=0}^{n} P_x(j) z^j = \prod_i (1 - x_i + x_i z).$$

At fixed n, t, the identity

$$\prod_i (1 + (x_i - \theta)z) = (1 - \theta z)^n F_{P_x}\left(1 + \frac{z}{1 - \theta z}\right) \tag{79}$$

shows that every $e_k(u)$ is a linear functional of P_x. To verify the identity, multiply each factor on the right by its factor $1 - \theta z$. The resulting factor is $1 + (x_i - \theta)z$. The right side is therefore a polynomial despite its rational notation. Also, $\sum_j v_j = 0$ gives

$$\kappa_y = -\frac{2e_2(v)}{n\sigma}. \tag{80}$$

Thus κ_y is linear in the second probability law at fixed mean.

Consider the contraction deficit $E(P_x, P_y) = Q(x, y) - S(x, y)$. For fixed P_y, it is a nonnegative weighted sum of squares of linear functionals of P_x, minus a linear functional. It is therefore convex in P_x. For fixed P_x, it is one plus a sum of nonnegative multiples of κ_y^k, $k \geq 2$, minus a linear functional of P_y. The powers are convex for $\kappa_y \geq 0$, so it is convex in P_y as well. These conclusions remain valid on the convex hull of the fixed-mean laws, because (80) stays nonnegative there. Only separate convexity is needed.

Lemma 8.2 (Convex fixed-mean extremizer)**.** *Fix $t \in [0, n]$. A continuous convex functional of a Poisson-binomial law with mean t attains its maximum at parameters whose values in $(0, 1)$ are all equal. Here convexity means convexity on a set containing the convex hull of the admissible laws.*

Proof. The parameter set $\{x \in [0, 1]^n : \sum_i x_i = t\}$ is compact. The law depends continuously on the parameters, so a maximum exists. Among all maximizers, choose one that maximizes $\sum_i x_i^2$. Suppose two unequal parameters p, q lie in $(0, 1)$. Keep $s = p + q$ fixed and vary $z = pq$. Their generating factor is

$$(1 - p + pw)(1 - q + qw) = (1 - s) + sw + z(1 - w)^2.$$

The full law is affine in z. Its feasible interval is $\max(0, s - 1) \leq z \leq s^2/4$. The current z lies strictly between the endpoints. The upper endpoint makes the parameters equal, and the lower endpoint makes one parameter zero or one.

The objective is convex in z. Write the current point as a strict convex combination of the two endpoints. Both endpoint values are at most the global maximum, while convexity bounds the current maximum by their weighted average. Both endpoint values must therefore equal the maximum. Moving to the lower endpoint preserves the objective and strictly increases $p^2 + q^2 = s^2 - 2z$. This contradicts the secondary maximization. Thus all interior parameters are equal. □

The linear case is the classical fixed-mean reduction used in the four-block argument [12]. The proof above gives the convex version required here. Maximize E over the two compact parameter sets at fixed n, t. Apply Lemma 8.2 to one law and then to the other. Each replacement preserves the global maximum. The maximum is therefore attained by two profiles of the form

$$\begin{aligned} x &= (1^a, q^m, 0^{n-a-m}), \quad q = \frac{t - a}{m}, \\ y &= (1^c, r^\ell, 0^{n-c-\ell}), \qquad r = \frac{t - c}{\ell}. \end{aligned} \tag{81}$$

An exponent here specifies the number of repeated entries. The integer indices satisfy

$$a, c \geq 0, \quad m, \ell \geq 1, \quad a + m, c + \ell \leq n, \quad L := \max(a, c) \leq t \leq R := \min(a + m, c + \ell). \tag{82}$$

Deterministic laws are represented by endpoints, as shown next.

8.3 Complete coverage of the profiles and endpoints

The profile $(a, m) = (0, n)$ makes $x_i = t/n$ for every i. It gives $u = 0$ and $A = J_n$ regardless of y. The same holds for the corresponding profile in the second slot. Those profiles can be omitted because $S = Q = 1$.

A deterministic law with mean $h \in \{1, \dots, n-1\}$ has h parameters equal to one and the rest zero. It is the endpoint $t = h$ of the profile $(a, m) = (h, 1)$. A nondeterministic partner has a support interval with h in its interior. Its integer upper endpoint is at least $h+1$, so its overlap with $[h, h+1]$ has positive length. Two deterministic laws with the same mean can both use $(h, 1)$. Thus every required pair has a representation with $L < R$ or is already a uniform case.

There is also a direct check for any discarded singleton overlap $L = R$. Both profiles then occur at endpoints, so both parameter vectors are binary with L ones. Up to row and column permutations, their matrix is $\operatorname{diag}(J_L, J_{n-L})$. Its Gram matrix is itself, so $S = Q$. An overlap with $L > R$ is infeasible.

For $L < R$, the interval endpoints with $0 < t < n$ are included by the same formulas or continuity. The original representation of a nonuniform matrix already has $0 < t < n$. The values $t = 0, n$ therefore omit no such matrix. For example, as $t \to 0$, each $x_i, y_j \le t$, so the first term of (75) tends to zero and the second tends to $1/n$. The limit at n follows by complementing the parameters.

Permuting x or y independently permutes rows or columns of A. Both $\operatorname{per}(A)$ and $\operatorname{per}(AA^{\mathsf{T}})$ are invariant under these operations. The certificate retains every ordered pair of profiles. It does not assume symmetry of the contraction under exchange of x and y.

8.4 Homogeneous integer certificates

Fix an admissible pair with $L < R$. For $U, V \ge 0$ with $U + V > 0$, put

$$T = LU + RV, \quad W = (n-L)U + (n-R)V, \quad D = TW, \quad t = \frac{T}{U+V},$$

and define

$$K = \ell D - n\big((L-c)U + (R-c)V\big)\big((\ell+c-L)U + (\ell+c-R)V\big). \tag{83}$$

For $0 < t < n$, $D > 0$ and

$$\sigma = \frac{D}{n^2(U+V)^2}, \qquad \kappa_y = \frac{K}{\ell D}.$$

The second identity follows from $\sum_j y_j(1-y_j) = \ell r(1-r) = (t-c)(\ell+c-t)/\ell$ and $\sum_j v_j^2 = n\sigma - \sum_j y_j(1-y_j)$.

Let X_k be the elementary symmetric polynomial of degree k in the following n linear forms:

$$\begin{array}{ll} mW, & a \text{ copies}, \\ (n(L-a) - mL)U + (n(R-a) - mR)V, & m \text{ copies}, \\ -mT, & n-a-m \text{ copies}. \end{array} \tag{84}$$

They are precisely the quantities $nm(U+V)(x_i - \theta)$. Define Y_k by replacing (a, m) with (c, ℓ). Thus

$$e_k(u) = \frac{X_k}{(nm(U+V))^k}, \qquad e_k(v) = \frac{Y_k}{(n\ell(U+V))^k}.$$

Set

$$\mathcal{C} = \sum_{k=0}^{n} k!(n-k)!(m\ell)^{n-k} X_k Y_k D^{n-k}, \tag{85}$$

$$\mathcal{G} = \sum_{k=0}^{n} k!(n-k)! m^{2(n-k)} \ell^{n-k} X_k^2 K^k D^{2(n-k)}. \tag{86}$$

Every coefficient is an integer. Substitution into (77) and (78) gives

$$S = \frac{\mathcal{C}}{n! m^n \ell^n D^n}, \qquad Q = \frac{\mathcal{G}}{n! m^{2n} \ell^n D^{2n}}. \tag{87}$$

Consequently,

$$S - Q = \frac{\mathcal{H}}{n! m^{2n} \ell^n D^{2n}}, \qquad \mathcal{H} := m^n \mathcal{C} D^n - \mathcal{G}, \tag{88}$$

$$S^2 - Q = \frac{\mathcal{S}}{(n!)^2 m^{2n} \ell^{2n} D^{2n}}, \qquad \mathcal{S} := \mathcal{C}^2 - n! \ell^n \mathcal{G}. \tag{89}$$

The denominators are positive on the relevant parameter range. The polynomial $\mathcal{C}$ has degree $2n$. Both deficit numerators have degree $4n$. Nonnegative coefficients of $U^{4n-j}V^j$ therefore prove their nonnegativity for all $U, V \geq 0$. This proves inequalities on whole mean intervals. The same signs are the Bernstein coefficient signs on $[L, R]$, because the basis conversion divides only by positive binomial coefficients.

8.5 The exhaustive finite calculation

The program `rank2_bernstein.py` implements (84)–(89) with arbitrary-precision Python integers. For a homogeneous polynomial of degree d, array position j stores the coefficient of $U^{d-j}V^j$. All $d+1$ positions are retained, including zeros. Addition is entrywise and multiplication is coefficient convolution. For linear forms $b_1, \ldots, b_n$, the elementary symmetric coefficients are constructed from

$$e_0^{(0)} = 1, \qquad e_k^{(0)} = 0 \ (k > 0), \qquad e_k^{(i)} = e_k^{(i-1)} + b_i e_{k-1}^{(i-1)}.$$

The implementation updates k in decreasing order, so every right-side value belongs to the previous stage.

For each order it uses exactly the profile list

$$\mathcal{P}_n = \{(a, m) : 0 \leq a < n, \ 1 \leq m \leq n - a\} \setminus \{(0, n)\}.$$

It checks every ordered pair in $\mathcal{P}_n^2$ with $L < R$. For each pair it constructs both degree-$4n$ numerator arrays and checks every coefficient. It records negative-coefficient failures, zero counts, and a SHA-256 digest of the complete ordered coefficient stream. It stops with an error if any negative coefficient occurs. The exhaustive mode uses no sampling flag.

The exhaustive computation gives the following counts. The coefficient column includes both numerator arrays.

Order	Ordered profile intervals	Coefficient positions
2	2	36
3	15	390
4	51	1 734
5	126	5 292
6	260	13 000
7	477	27 666
8	805	53 130
9	1 276	94 424
10	1 926	157 932
11	2 795	251 550
12	3 927	384 846
13	5 370	569 220
14	7 176	818 064
15	9 401	1 146 922
Total	33 607	3 524 206

Every coefficient is nonnegative. The certificate therefore proves $S - Q \geq 0$ on each enumerated interval. The representation, convex reduction, and coverage argument extend this result to every rank-two doubly stochastic matrix in the stated orders. The independently checked $\mathcal{S}$ arrays also verify the strong inequality on each profile interval. The main theorem needs only the contraction certificate.

A separate rational implementation starts from polynomials in the mean t and then converts to Bernstein coefficients. It agrees coefficient by coefficient with the integer construction on all 454 profile intervals in orders two through six. This compares 20 452 positions exactly. A direct subset-permanent computation checks 378 profile pairs in orders two through seven. It gives 1 390 exact identity comparisons, 21 deterministic cases, and 251 uniform-profile instances. A second direct computation checks another 209 rational instances, including valid endpoints. These finite checks test the identities and implementation. The universal parameter coverage comes from the preceding analytic proof and the exhaustive coefficient signs.

The four entry points use the Python standard library:

```
python3 rank2_bernstein.py --first 2 --last 15
python3 rank2_bernstein_fraction_check.py --first 2 --last 6
python3 rank2_identity_audit.py
python3 rank2_boundary_audit.py
```

They are included with the accompanying source and verification records. The first two JSON reports contain timing fields, which can change on a repeated run. The coefficient digests, counts, and mathematical results are the reproducibility criteria.

8.6 Conclusion and equality

Proof of Theorem 8.1. The certificate and analytic reduction give $\operatorname{per}(AA^{\mathsf{T}}) \leq \operatorname{per}(A)$ for every matrix in the stated range. Suppose now that $A \neq J_n$. Both centered vectors in (76) are then nonzero. Consequently $\kappa_y > 0$, and

$$e_2(u) = -\frac{1}{2}\sum_i u_i^2 < 0.$$

The $k = 2$ term of (78) is strictly positive. All its other terms are nonnegative, and its zeroth term is one. Hence

$$d \geq \operatorname{per}(AA^{\mathsf{T}}) > c_n.$$

Transposition preserves the rank condition, so the same contraction applies to A^{T}. Lemma 2.2 then gives the third contraction in (73), in the original indexing. Their weighted sum gives

$$\Delta_n(A) \geq 4d(d - c_n) > 0.$$

The strong inequality is also strict, since

$$d^2 \geq d \operatorname{per}(AA^{\mathsf{T}}) > c_n \operatorname{per}(AA^{\mathsf{T}}).$$

At $A = J_n$, every product matrix in the statement is J_n and every corresponding permanent is c_n. Thus equality holds in both inequalities precisely in this case. □

All row and column types are included in these orders. The result does not remove the order cutoff for arbitrary rank-two matrices. A proof of the coefficient signs uniformly in n, or a different all-order argument, remains necessary.

9 Closure, sparse perturbations, and boundary reduction

Direct sums and row and column permutations extend the matrix families above. The sparse-perturbation theorems give additional noncirculant families and explicit margins. The final boundary theorem is a reduction of the unrestricted strong inequality. It does not prove the required boundary inequality.

9.1 Direct sums and independent permutations

Lemma 9.1 (Closure of the strong bounds)**.** *Let $B_i \in \Omega_{m_i}$ for $1 \leq i \leq r$, with $m_i \geq 1$ and $\sum_i m_i = n$. Suppose each B_i satisfies both strong bounds* (2) *in its own dimension. Then their direct sum $B = \bigoplus_i B_i$ satisfies both strong bounds in dimension n. The same holds for $A = UBV$, where U, V are arbitrary permutation matrices. Each such A satisfies Gyires's inequality. If $r \geq 2$, both strong bounds and Gyires's inequality are strict.*

Proof. Permanents multiply over direct sums. Their two Gram matrices also decompose into the corresponding direct sums. Thus

$$\operatorname{per}(B)^2 \geq \left(\prod_i c_{m_i}\right) \operatorname{per}(BB^{\mathsf{T}}), \qquad \operatorname{per}(B)^2 \geq \left(\prod_i c_{m_i}\right) \operatorname{per}(B^{\mathsf{T}}B).$$

The multinomial theorem gives

$$1 = \left(\sum_i \frac{m_i}{n}\right)^n \geq \frac{n!}{\prod_i m_i!} \prod_i \left(\frac{m_i}{n}\right)^{m_i} = \frac{c_n}{\prod_i c_{m_i}}.$$

For $r \geq 2$, there are other positive terms in that multinomial expansion. The displayed inequality is then strict. Both Gram permanents are positive because their positive diagonal entries supply a positive permanent term. This proves the strong bounds and their strictness.

For $A = UBV$, one has

$$AA^{\mathsf{T}} = U(BB^{\mathsf{T}})U^{\mathsf{T}}, \qquad A^{\mathsf{T}}A = V^{\mathsf{T}}(B^{\mathsf{T}}B)V, \qquad \operatorname{per}(A) = \operatorname{per}(B).$$

Permutation invariance therefore preserves both strong bounds. Lemma 2.3 now gives Gyires's inequality, with strictness when both bounds are strict. This last step applies to the original indexing of A^2. □

This lemma permits different admissible matrices in different blocks. For example, different permutation cycles can use different nonnegative triples of weights after stochastic normalization. It also applies to blocks from the other proved matrix families.

9.2 Positive-semidefinite matrices

A real symmetric matrix is positive semidefinite if $x^{\mathsf{T}}Ax \geq 0$ for every real vector x. The next proposition follows from monotonicity in positive-semidefinite order and the van der Waerden bound.

Proposition 9.2. *If $A \in \Omega_n$ is symmetric and positive semidefinite, then it satisfies both strong bounds and Gyires's inequality. Equality in any of these inequalities holds only at J_n. The conclusion is preserved by independent row and column permutations.*

Proof. Every eigenvalue of A lies in $[0, 1]$. Nonnegativity and row sums one bound the spectral radius by one, while positive semidefiniteness gives the lower bound. Hence $0 \preceq A^2 \preceq A$. The permanent is increasing in positive-semidefinite order. To see this, if $0 \preceq X \preceq Y$, a telescoping expansion gives $X^{\otimes n} \preceq Y^{\otimes n}$. Each term in the difference is a tensor product of positive-semidefinite matrices. For

$$v = \frac{1}{\sqrt{n!}} \sum_{\sigma \in S_n} e_{\sigma(1)} \otimes \cdots \otimes e_{\sigma(n)},$$

direct expansion gives $v^{\mathsf{T}} X^{\otimes n} v = \operatorname{per}(X)$. Thus $\operatorname{per}(A^2) \leq \operatorname{per}(A)$. The van der Waerden bound and its equality case give

$$c_n \operatorname{per}(A^2) \leq c_n \operatorname{per}(A) \leq \operatorname{per}(A)^2,$$

with the second inequality strict when $A \neq J_n$. Since both Gram matrices equal A^2, this proves all assertions for A. The permutation formulas in Lemma 9.1 preserve the Gram bounds. Lemma 2.3 gives the final conclusion after relabeling. □

9.3 Three exceptional rows and an exact cubic expansion

A row is called exceptional if it differs from the corresponding row of J_n. For $n \geq 5$, define

$$K_n = \frac{6(n-5)^4 + 77(n-5)^3 + 333(n-5)^2 + 515(n-5) + 149}{6(n-1)^2(n-2)^2} > 0. \tag{90}$$

The squared Frobenius norm is $\|B\|_F^2 = \sum_{i,j} b_{ij}^2$.

Theorem 9.3 (Three exceptional rows)**.** *Let $A \in \Omega_n$, with $n \geq 5$, and suppose at most three rows are exceptional. Then*

$$\operatorname{per}(A)^2 - c_n \operatorname{per}(AA^{\mathsf{T}}) \geq c_n^2 K_n \|A - J_n\|_F^2. \tag{91}$$

Equality in the corresponding strong inequality occurs only at J_n. If $A - J_n$ is supported on at most three rows and at most three columns, then

$$\Delta_n(A) \geq 4c_n^2 K_n \|A - J_n\|_F^2. \tag{92}$$

Proof. Permute rows so that the first three contain all exceptions. Set $B = A - J_n$, $s = \sum b_{ij}^2$, and $t = \sum b_{ij}^3$. Both row sums and column sums of B vanish. Every entry satisfies $b_{ij} \geq -1/n$. For a zero-sum matrix B, let $M_k(B)$ be the sum of the permanents of all its $k \times k$ submatrices, with $M_0 = 1$. Expansion by the entries taken from B gives

$$\operatorname{per}(J_n + B) = \sum_{k=0}^{n} (n-k)! n^{-(n-k)} M_k(B). \tag{93}$$

The zero sums imply

$$M_1(B) = 0, \qquad M_2(B) = \frac{1}{2} \sum b_{ij}^2, \qquad M_3(B) = \frac{2}{3} \sum b_{ij}^3.$$

Here $k!M_k$ is the sum over ordered distinct row indices and ordered distinct column indices. Apply inclusion-exclusion to these distinctness conditions. A row or column index that occurs only once makes the unrestricted sum zero. At $k = 2$, only the identification of both row indices and both column indices survives. At $k = 3$, only the identification of all three row indices and all three column indices survives. The two inclusion-exclusion factors in that case are both 2. Division by 2! or 3! gives the stated identities. Since only three rows of B can be nonzero, every $M_k(B)$ with $k > 3$ vanishes. Put

$$\alpha = \frac{n}{2(n-1)}, \qquad \beta = \frac{2n^2}{3(n-1)(n-2)}.$$

Equation (93) becomes

$$\frac{\operatorname{per}(A)}{c_n} = 1 + \alpha s + \beta t. \tag{94}$$

In each column, the three potentially nonzero deviations x, y, z have sum zero and are at least $-1/n$. They satisfy

$$x^3 + y^3 + z^3 \geq -\frac{x^2 + y^2 + z^2}{2n}.$$

If the cubic sum is negative, exactly one entry is negative. Write it as $-r$, with $0 < r \leq 1/n$, and the others as $u, v \geq 0$, where $u + v = r$. The cubic sum is $-3ruv$. Also $r^2 + u^2 + v^2 = 2r^2 - 2uv \geq 6uv$, since $r^2 \geq 4uv$. This proves the inequality; a nonnegative cubic sum needs no estimate. Summing over columns gives $t \geq -s/(2n)$. Consequently,

$$\frac{\operatorname{per}(A)}{c_n} \geq 1 + as, \qquad a = \alpha - \frac{\beta}{2n} = \frac{n(3n-8)}{6(n-1)(n-2)} > 0. \tag{95}$$

The three exceptional entries of each column of A sum to $3/n$. Their squared deviations therefore sum to at most $(3/n)^2 - 3/n^2 = 6/n^2$. Thus $0 \leq s \leq 6/n$.

Let $D = BB^{\mathsf{T}}$. It is positive semidefinite, has zero row sums, and is supported on three rows and three columns. Its trace is s, and $AA^{\mathsf{T}} = J_n + D$. Write $E = \sum d_{ij}^2$ and $U = \sum d_{ij}^3$. The nonnegative eigenvalues give $E \leq s^2$. Also,

$$0 \leq U \leq s^3/6. \tag{96}$$

For this estimate, use the plane $H = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\}$. For $x \in H$, the restriction of $\operatorname{diag}(x)$ to the orthonormal basis $(1, -1, 0)/\sqrt{2}$, $(1, 1, -2)/\sqrt{6}$ is

$$\begin{pmatrix} -x_3/2 & (x_1 - x_2)/(2\sqrt{3}) \\ (x_1 - x_2)/(2\sqrt{3}) & x_3/2 \end{pmatrix}.$$

Its eigenvalues are $\pm\|x\|/\sqrt{6}$. Hence for $x, y, z \in H$,

$$\left|\sum_i x_i y_i z_i\right| \le \frac{1}{\sqrt{6}}\|x\|\|y\|\|z\|.$$

Take a spectral decomposition $D = \sum_r \lambda_r u_r u_r^{\mathsf{T}}$ with $\lambda_r \ge 0$ and unit $u_r \in H$. Then

$$U = \sum_{r,s,t} \lambda_r\lambda_s\lambda_t \left(\sum_i u_{r,i}u_{s,i}u_{t,i}\right)^2 \le \frac{1}{6}\left(\sum_r \lambda_r\right)^3.$$

This proves (96). Apply the same cubic permanent expansion to D to obtain

$$\frac{\operatorname{per}(AA^{\mathsf{T}})}{c_n} = 1 + \alpha E + \beta U \le 1 + \alpha s^2 + \frac{\beta}{6}s^3.$$

Together with (95), this yields

$$\frac{\operatorname{per}(A)^2 - c_n \operatorname{per}(AA^{\mathsf{T}})}{c_n^2} \ge sF_n(s), \qquad F_n(s) = 2a + (a^2 - \alpha)s - \frac{\beta}{6}s^2.$$

One has $a \le 2/3$ and $\alpha \ge 1/2$, so $a^2 - \alpha < 0$. Thus F_n decreases for $s \ge 0$. Direct algebra gives

$$F_n(6/n) = \frac{6n^4 - 43n^3 + 78n^2 - 40n + 24}{6(n-1)^2(n-2)^2} = K_n.$$

The expansion in (90) proves $K_n > 0$ for $n \ge 5$. This proves (91) and its strictness away from J_n.

Under the additional column-support condition, apply the same theorem to A^{T}. Both Gram permanents are then at most $\operatorname{per}(A)^2/c_n - c_n K_n s$. Lemma 2.3 gives $2\operatorname{per}(A^2) \le \operatorname{per}(AA^{\mathsf{T}}) + \operatorname{per}(A^{\mathsf{T}}A)$. Adding the two estimates proves (92). □

For example, replace any chosen 3×3 block of J_n by $(3/n)S$, where $S \in \Omega_3$. The resulting matrix satisfies the support condition in (92). The chosen row set and column set need not agree. The family includes zero entries. The one-sided statement (91) alone does not imply Gyires's inequality without a bound for the other Gram matrix.

9.4 A trace bound for centered positive-semidefinite perturbations

The next estimate controls that other Gram matrix when the order is sufficiently large. It imposes no restriction on the number of nonzero rows of its input.

Lemma 9.4. *Let D be a real positive-semidefinite $n \times n$ matrix, with $n \ge 3$, zero row sums, and trace $s < 1/e$. Then*

$$\frac{\operatorname{per}(J_n + D)}{c_n} \le 1 + \frac{n}{2(n-1)}s^2 + \frac{2n^2}{3(n-1)(n-2)}s^3 + \frac{(es)^4}{1-es}. \tag{97}$$

Neither D nor $J_n + D$ is required to have nonnegative entries.

Proof. Let Y be a centered circular complex Gaussian vector with covariance D. For example, take independent standard circular complex Gaussian coordinates and apply $D^{1/2}$. The expectation of the squared absolute value of $\sum_i Y_i$ is $\mathbf{1}^{\mathsf{T}}D\mathbf{1} = 0$. Thus $\sum_i Y_i = 0$ almost surely.

Let Z be an independent standard circular complex Gaussian variable. Gaussian moment pairing gives

$$\mathbb{E}\prod_i |W_i|^2 = \operatorname{per}(\mathbb{E}[WW^*]).$$

Indeed, after expansion in independent standard coordinates, a mixed moment is zero unless the two coordinate multiplicities agree. When they agree with multiplicities α, the moment is $\alpha!$. This is the same counting identity as Lemma 2.1. Apply it to $W_i = Z/\sqrt{n} + Y_i$ and expand the product. Rotational symmetry eliminates terms with different powers of Z. Since $\mathbb{E}|Z|^{2j} = j!$, one obtains

$$\frac{\operatorname{per}(J_n + D)}{c_n} = \sum_{k=0}^{n} \frac{n^k}{(n)_k} \mathbb{E}|e_k(Y)|^2, \qquad (n)_k = n(n-1)\cdots(n-k+1). \tag{98}$$

Here e_k is the elementary symmetric polynomial of degree k. The zeroth term is one. The first term is zero.

If $\sum_i y_i = 0$, the elementary inequality $|1 + w| \le \exp(\operatorname{Re} w + |w|^2/2)$ gives

$$\left|\prod_i (1 + z y_i)\right| \le \exp(|z|^2 \|y\|^2/2).$$

Cauchy's coefficient estimate on a circle of radius r, optimized at $r^2 = k/\|y\|^2$, gives $|e_k(y)|^2 \le (e/k)^k \|y\|^{2k}$ for $k \ge 1$. The zero vector satisfies this directly. If $\lambda_i \ge 0$ are the eigenvalues of D, Gaussian moments give

$$\mathbb{E}\|Y\|^{2k} = k! \sum_{|\alpha|=k} \prod_i \lambda_i^{\alpha_i} \le k! \left(\sum_i \lambda_i\right)^k = k! s^k.$$

Thus the kth term of (98) is at most

$$\frac{n^k}{(n)_k} \left(\frac{e}{k}\right)^k k! s^k = \frac{(ens/k)^k}{\binom{n}{k}} \le (es)^k.$$

The last inequality uses $\binom{n}{k} = \prod_{i=0}^{k-1} (n-i)/(k-i) \ge (n/k)^k$.

The exact low-degree identities from (93) give the second term $\alpha \sum d_{ij}^2$ and third term $\beta \sum d_{ij}^3$, with the same α, β as above. Positive semidefiniteness gives $\sum d_{ij}^2 \le s^2$ and $|d_{ij}| \le \sqrt{d_{ii} d_{jj}} \le s$. Consequently $\sum d_{ij}^3 \le \sum |d_{ij}|^3 \le s^3$. Use these two sharper bounds and sum the geometric tail from degree four. This proves (97). □

Theorem 9.5 (Three exceptional rows in orders at least 31)**.** *Let $A \in \Omega_n$ with $n \ge 31$. If at most three rows are exceptional, then, for $\gamma = 9301/445904 > 0$,*

$$\operatorname{per}(A)^2 - c_n \operatorname{per}(AA^{\mathsf{T}}) \ge c_n^2 \gamma \|A - J_n\|_F^2, \qquad \operatorname{per}(A)^2 - c_n \operatorname{per}(A^{\mathsf{T}}A) \ge c_n^2 \gamma \|A - J_n\|_F^2.$$

Hence $\Delta_n(A) \ge 4 c_n^2 \gamma \|A - J_n\|_F^2$. Equality in either strong inequality or in Gyires's inequality occurs only at J_n. The same conclusions hold if at most three columns are exceptional.

Proof. Use B, s, a, α, β from Theorem 9.3. Then $\operatorname{per}(A)/c_n \ge 1 + as$, $s \le 6/n$, and $a \ge 1/2$ for $n \ge 6$. Both BB^{T} and $B^{\mathsf{T}}B$ are positive semidefinite, with zero row sums and trace s. For $n \ge 31$, one has $s \le 6/31 < 1/e$. Lemma 9.4 therefore applies to both Gram matrices. Also $a^2 - \alpha \ge -1/3$ and $\beta \le 3/4$. The first follows from $a^2 \ge 1/4$ and $\alpha \le 31/60 < 7/12$. The second is equivalent to $n^2 - 27n + 18 \ge 0$. For either Gram matrix G, these bounds give

$$\frac{\operatorname{per}(A)^2 - c_n \operatorname{per}(G)}{c_n^2} \ge s\left(1 - \frac{s}{3} - \frac{3s^2}{4} - \frac{e^4 s^3}{1 - es}\right).$$

Every subtracted term increases with s on $[0,1/e)$. Use $s \le 6/31$ and $e < 11/4$ to obtain

$$1 - \frac{2}{31} - \frac{27}{961} - \frac{(11/4)^4(6/31)^3}{1-(11/4)(6/31)} = \frac{9301}{445904}.$$

For the elementary exponential estimate, the series for e gives $e < 8/3 + (1/24)\sum_{j\ge 0} 5^{-j} = 87/32 < 11/4$. The two Gram bounds follow. Lemma 2.3 gives the Gyires margin. Transposition proves the column version. □

The threshold 31 is sufficient for these estimates. It is not claimed to be optimal. For orders 5 through 30, Theorem 9.3 gives the full Gyires conclusion only with its additional three-column support condition.

9.5 The symmetrized denominator along a radial segment

For a matrix with all row sums and column sums one, define

$$\mathcal{Q}(A) = \frac{\operatorname{per}(AA^{\mathsf T}) + \operatorname{per}(A^{\mathsf T}A) + 2\operatorname{per}(A^2)}{4}, \qquad r_A = \frac{p_A + p_{A^{\mathsf T}}}{2}.$$

Lemma 2.1 gives

$$\mathcal{Q}(A) = \|r_A\|_F^2, \qquad [z_1 \cdots z_n] r_A = \operatorname{per}(A).$$

One quarter of Gyires's denominator is therefore the squared Fischer norm of the average of the row-product and column-product polynomials.

Proposition 9.6. *Let $n \ge 2$, $B\mathbf{1} = B^{\mathsf T}\mathbf{1} = 0$, and $A_t = J_n + tB$. There are numbers $q_k \ge 0$ such that*

$$\mathcal{Q}(A_t) = c_n + \sum_{k=2}^{n} q_k t^{2k}.$$

Moreover,

$$q_2 = \frac{(n-2)!}{n^{n-2}} \frac{\operatorname{tr}[(B^{\mathsf T}B + BB^{\mathsf T})^2]}{8}.$$

If $B \ne 0$, then $q_2 > 0$ and $\mathcal{Q}(A_t)$ is strictly increasing for $t > 0$. The identity does not require nonnegative entries in A_t.

Proof. Use orthonormal coordinates $z = z_0\mathbf{1}/\sqrt{n} + y$ with $y \perp \mathbf{1}$. Each row form of A_t is $z_0/\sqrt{n} + t(By)_i$. Set

$$r_k(y) = \frac{e_k(By) + e_k(B^{\mathsf T}y)}{2}.$$

Then $r_0 = 1$, $r_1 = 0$, and

$$r_{A_t}(z) = \sum_{k=0}^{n} t^k \frac{z_0^{n-k}}{n^{(n-k)/2}} r_k(y).$$

The factorial inner product is invariant under orthogonal coordinate changes. One direct justification is its Gaussian representation $\|p\|_F^2 = \mathbb{E}|p(Z)|^2$ for independent standard circular complex Gaussian coordinates Z. The moment calculation in Lemma 9.4 proves this representation, and orthogonal changes preserve the Gaussian law. Different powers of z_0 are orthogonal. Therefore

$$q_k = \frac{(n-k)!}{n^{n-k}} \|r_k\|_F^2 \ge 0.$$

With $M = B^{\mathsf T}B + BB^{\mathsf T}$, the zero-sum identities give $r_2(y) = -y^{\mathsf T}My/4$. For a real symmetric matrix, direct coefficient expansion gives $\|y^{\mathsf T}My\|_F^2 = 2\operatorname{tr}(M^2)$. The matrix M annihilates $\mathbf{1}$,

so its trace on the contrast space is its full trace. This proves the formula for q_2. Finally, M is positive semidefinite and is zero exactly when $B = 0$. □

The corresponding ratio need not increase. This limitation already occurs for a doubly stochastic circulant of order three. Let P be a three-cycle, let $A = (I + P)/2$, and set $A_t = J_3 + t(A - J_3)$ for $0 \le t \le 1$. Its circulant weights are (u, u, v), where $u = 1/3 + t/6$ and $v = 1/3 - t/3$. Transposition permutes the row forms, so $\mathcal{Q}(A_t) = \operatorname{per}(A_t A_t^{\mathsf{T}})$. Direct permanent expansion gives

$$p(t) = \operatorname{per}(A_t) = \frac{2}{9} + \frac{t^2}{12} - \frac{t^3}{18}, \qquad q(t) = \mathcal{Q}(A_t) = \frac{2}{9} + \frac{t^4}{48} + \frac{t^6}{144}.$$

Thus $p(1) = q(1) = 1/4$, $p'(1) = 0$, and $q'(1) = 1/8$. It follows that

$$\frac{d}{dt}\frac{p(t)^2}{q(t)}\bigg|_{t=1} = -\frac{1}{8} < 0.$$

The endpoint ratio is $1/4 > c_3 = 2/9$. The example therefore refutes radial monotonicity but does not refute Gyires's inequality.

9.6 A direct reduction to the matrix boundary

The boundary of Ω_n consists of its matrices with at least one zero entry. For a row, its support is the set of columns in which its entries are positive.

Theorem 9.7 (Boundary reduction)**.** *Fix $n \ge 2$. The first strong bound* (2) *holds on all of Ω_n if and only if it holds on the boundary. If it holds on the boundary, it is strict at every strictly positive nonuniform matrix. More generally, every global maximizer on Ω_n of*

$$F(A) = \sqrt{c_n}\,\|p_A\|_F - \operatorname{per}(A)$$

has identical rows within each row-support class.

Proof. The function F is continuous on the compact set Ω_n, so it has a global maximizer. Suppose a matrix has two distinct rows r, s with the same support. Put $\mu = (r + s)/2$, $d = (r - s)/2$, and

$$T = \min_{j : d_j \neq 0} \frac{\mu_j}{|d_j|}.$$

The minimum is over a nonempty finite set. On the common support, both entries are positive, so $\mu_j > |d_j|$ when $d_j \neq 0$. Hence $T > 1$. Replace the two rows by $\mu + td$ and $\mu - td$ for $0 \le t \le T$. The resulting matrix $A(t)$ remains doubly stochastic. Its row sums stay one, its column sums do not change, and the definition of T gives nonnegative entries. The original matrix is $A(1)$, and $A(T)$ has a zero entry. If $R(z)$ is the product of the unchanged row forms, then

$$p_{A(t)} = \left[(\mu \cdot z)^2 - t^2 (d \cdot z)^2\right] R(z).$$

This polynomial is affine in t^2. The two endpoint polynomials are distinct because $d \neq 0$ and R is nonzero. Both have value one at $\mathbf{1}$, so they cannot be positive scalar multiples. Strict convexity of a Hilbert norm on the segment between such vectors gives

$$F(A) < (1 - T^{-2})F(A(0)) + T^{-2}F(A(T)).$$

The subtracted permanent is a linear coefficient functional, so it does not affect this strict inequality. A global maximizer cannot have such a row pair. This proves the support-class assertion.

If a global maximizer is strictly positive, every row has the same support. All rows must then agree, and the column sums force that matrix to be J_n. Since $F(J_n) = 0$, it follows that

$$\max_{A\in\Omega_n} F(A) = \max\left\{0, \max_{A\in\partial\Omega_n} F(A)\right\}.$$

The first strong inequality is equivalent to $F(A) \le 0$ because both terms before squaring are nonnegative. This proves the boundary equivalence. If boundary validity holds, the entire set has $F \le 0$. Every strictly positive nonuniform matrix has two distinct rows with common support. The strict interpolation estimate therefore gives $F(A) < 0$ there. □

The boundary has the same order as the original matrix. The theorem is not a reduction to permutation matrices and is not an induction on the dimension. No boundary-validity assertion for arbitrary order is assumed in the preceding matrix-family proofs.

10 Scalar and two-group polynomial inequalities

Products of nonnegative linear forms connect the matrix inequalities to coefficient norm bounds. The scalar inequalities in this section also apply to a larger class of polynomials. The scalar bounds precede their polynomial interpretations.

For a homogeneous polynomial p of degree n in n variables, write $N(p) = \langle p, p\rangle_F$ for the squared Fischer norm from Section 2. Call p normalized when $p(\mathbf{1}) = 1$. It is balanced when $\partial_i p(\mathbf{1}) = 1$ for every variable. Its nonnegative coefficients then define a probability law on exponent vectors, with mean one in each coordinate.

A nonnegative sequence has interval support if its positive terms occupy consecutive indices. It is log-concave if $u_k^2 \ge u_{k-1}u_{k+1}$ at every interior index. A likelihood ratio is the sequence P_k/B_k for probability laws P, B with every $B_k > 0$. The scalar assumptions below concern this ratio, not the reference law itself.

The term Lorentzian is used only with its stated coefficient criterion in this section. For a homogeneous binary polynomial $\sum_{k=0}^{n} a_k X^k Y^{n-k}$ with nonnegative coefficients, that criterion is interval support and log-concavity of $a_k/\binom{n}{k}$. Replacement of X, Y by sums of disjoint groups of variables gives the corresponding two-group class. These are standard Lorentzian closure and coefficient properties [3]. Each theorem states its coefficient assumptions explicitly.

10.1 Localization of likelihood ratios

Let $B = (B_0, \dots, B_n)$ be a strictly positive probability law, and let $m \in \{1, \dots, n-1\}$. Let $\mathcal{U}$ contain the nonnegative log-concave sequences with interval support that satisfy

$$u_m = 1, \qquad \sum_{k=0}^{n} (k-m) B_k u_k = 0.$$

The following localization statement will be used three times. Every continuous strictly convex function of the coordinates u attains its maximum on $\mathcal{U}$. Every maximizer is either the singleton $u_m = 1$, with all other entries zero, or has the form

$$u_k = t^{k-m}\mathbf{1}_{\{a\le k\le b\}}, \qquad a < m < b,$$

where $t > 0$ is the unique value that satisfies the mean equation. Translation of the indices gives the same statement for any finite integer interval and an interior integer mean.

To prove compactness, put $Z = \sum B_k u_k$ and $P_k = B_k u_k / Z$. This law has mean m. Concavity of the piecewise-linear interpolation of $\log u$ on its positive support gives

$$\sum_k P_k \log u_k \le \log u_m = 0.$$

The left side is $D(P\|B) + \log Z$, where $D(P\|B) = \sum_{P_k>0} P_k \log(P_k/B_k)$ is relative entropy. Jensen's inequality gives $D(P\|B) \ge 0$. Hence $Z \le 1$ and $0 \le u_k \le B_k^{-1}$. The feasible set is closed. Indeed, its multi-step log-concavity inequalities

$$u_j^{k-i} \ge u_i^{k-j} u_k^{j-i} \qquad (i < j < k)$$

are continuous polynomial inequalities. They exclude an internal zero between positive limiting endpoints. Thus $\mathcal{U}$ is compact.

If a support endpoint equals m, balance forces the singleton case. Otherwise write the support as $[a, b]$, with $a < m < b$. At each active log-concavity equality, impose

$$2h_k = h_{k-1} + h_{k+1}.$$

If any interior inequality is strict, these independent second-difference equations leave a solution space of dimension at least three. The two further equations

$$h_m = 0, \qquad \sum_{k=a}^{b} (k-m) B_k u_k h_k = 0$$

therefore leave a nonzero solution. For small $\varepsilon > 0$, put $u_k^{\pm} = u_k(1 \pm \varepsilon h_k)$ on the same support. Both constraints hold exactly. At an active equality, the new normalized gap is

$$(1 \pm \varepsilon h_k)^2 - (1 \pm \varepsilon h_{k-1})(1 \pm \varepsilon h_{k+1}) = \varepsilon^2 (h_k - h_{k-1})^2 \ge 0.$$

Strict inequalities persist for small ε. Thus u^+ and u^- are feasible and have average u. Strict convexity makes the average of their objective values larger than its value at u. This contradicts maximality. Every interior log-concavity inequality is therefore equality, which gives the geometric form. The mean of a positive exponential tilt is strictly increasing in $\log t$, because its derivative is the positive variance. Its limits are a and b. This proves the asserted existence and uniqueness of t.

10.2 The mean-one binomial theorem

Let $n \ge 2$, let B be the probability mass function of $\mathrm{Bin}(n, 1/n)$, and let P be a probability mass function on $\{0, \dots, n\}$ with mean one. Suppose

$$r_k = \frac{P_k}{B_k}$$

is log-concave, with interval support. Then

$$\boxed{\sum_{k=0}^{n} B_k r_k^2 \le r_1^2.} \tag{99}$$

Equality holds exactly when $P = B$.

The interval-support and mean assumptions imply $P_1 > 0$. Set $u_k = (P_k/B_k)/(P_1/B_1)$. Then $u_1 = 1$, and the mean equation is

$$\sum_{k=0}^{n}(k-1)B_k u_k = 0. \tag{100}$$

The assertion is $\Phi(u) = \sum B_k u_k^2 \le 1$. The preceding localization applies to this strictly convex objective. A singleton has objective $B_1 < 1$. Every other candidate has support $\{0, \dots, L\}$ with $L \ge 2$ and

$$u_k = t^{k-1}\ (0 \le k \le L), \qquad u_k = 0\ (k > L). \tag{101}$$

Put $F_L(z) = \sum_{k=0}^{L} B_k z^k$. The balancing root satisfies

$$tF_L'(t) = F_L(t), \quad \text{equivalently} \quad \sum_{k=2}^{L}(k-1)B_k t^k = B_0. \tag{102}$$

The left side is strictly increasing for $t > 0$. Its root is $t_n = 1$ for $L = n$ and $t_L > 1$ for a proper truncation. It remains to prove

$$\Phi(u) = F_L(t_L^2)/t_L^2 \le 1. \tag{103}$$

10.2.1 The balancing tilt

Put $c_k = B_k/B_0$. For $k \ge 1$,

$$c_k = \frac{\binom{n}{k}}{(n-1)^k} \le \frac{n}{n-1}\frac{1}{k!}. \tag{104}$$

Indeed the first factor in $(n)_k/(n-1)^k$ is $n/(n-1)$, the second is one, and all subsequent factors are at most one.

Let

$$G_L(t) = \sum_{k=2}^{L}(k-1)c_k t^k.$$

Then $G_L(t_L) = 1$, and

$$1 - G_L(1) = \sum_{k>L}(k-1)c_k \le \frac{n}{n-1}\sum_{k>L}\frac{k-1}{k!} = \frac{n}{n-1}\frac{1}{L!}.$$

For $t \ge 1$,

$$G_L'(t) \ge 2c_2 = \frac{n}{n-1}.$$

Consequently,

$$\boxed{0 \le t_L - 1 \le \frac{1}{L!}.} \tag{105}$$

This bound controls the change caused by truncating the binomial law.

10.2.2 Comparison with the omitted binomial tail

Fix $L < n$ and abbreviate $t = t_L > 1$. Let

$$M = \sum_{k>L} B_k, \qquad T = \sum_{k>L} (k-1)B_k.$$

Subtracting the mean equation at $t = 1$ from (102) gives

$$T = \sum_{k=2}^{L} (k-1)B_k(t^k - 1). \tag{106}$$

Put $f(z) = F_L(z)/z$. Eliminating B_0 by (102) gives the exact identity

$$f(t^2) - f(1) = \sum_{k=3}^{L} B_k \psi_k(t), \tag{107}$$

where

$$\psi_k(t) = t^{2k-2} - 1 - (k-1)t^{k-2}(t^2 - 1).$$

The $k = 2$ contribution vanishes exactly.

If $L = 2$, (107) is zero, and $f(t^2) = f(1) = 1 - M < 1$, which proves (103) in that case.

Now assume $L \geq 3$. Write $s = \log t$ and $r = k - 1$. Then

$$\psi_k(t) = 2e^{rs}\big(\sinh(rs) - r\sinh s\big).$$

Taylor's integral remainder and $\sinh s \geq s$ give

$$\sinh(rs) - r\sinh s \leq \sinh(rs) - rs \leq \frac{r^3 s^3}{6}\cosh(rs).$$

Also $e^{(r+1)s} - 1 \geq (r+1)s$. Hence

$$\frac{\psi_k(t)}{(k-1)(t^k-1)} \leq \frac{r^2}{3(r+1)} e^{2rs} s^2 \leq \frac{r}{3} e^{2rs} s^2.$$

By (105), $s \leq 1/L!$ and $r \leq L - 1$, so $2rs \leq 2/3$ and $e^{2rs} < 2$. This gives

$$\frac{\psi_k(t)}{(k-1)(t^k-1)} < C_L := \frac{2(L-1)}{3(L!)^2} < \frac{1}{L+1}. \tag{108}$$

The final inequality follows already from $L! \geq L$.

The binomial tail satisfies a complementary estimate. For $k \geq L + 1$,

$$\frac{B_{k+1}}{B_k} = \frac{n-k}{(n-1)(k+1)} \leq \frac{1}{L+2}.$$

Conditioned on $K \geq L + 1$, the excess $K - (L + 1)$ is stochastically bounded by a geometric distribution with successive probability ratio $\rho = 1/(L+2)$. More explicitly, write $b_i = B_{L+1+i}$,

with zero values past the end of the support. The coefficient-ratio bound gives $b_{i+r} \le \rho^r b_i$ term by term, and hence

$$\mathbb{P}(K-(L+1) \ge r \mid K > L) = \frac{\sum_{i\ge 0} b_{i+r}}{\sum_{i\ge 0} b_i} \le \rho^r.$$

Summing these tail probabilities gives

$$\frac{T}{M} = \mathbb{E}[K-1 \mid K > L] \le L + \frac{1}{L+1} < L+1. \tag{109}$$

Combining (106)–(109),

$$f(t^2) - f(1) < \frac{T}{L+1} < M.$$

Since $f(1) = F_L(1) = 1 - M$, this proves $f(t^2) < 1$. Together with the $L = 2$ case, every proper truncation has strictly smaller objective than one.

For $L = n$, the unique root is $t = 1$ and the objective is exactly one. The degenerate support $\{1\}$ also has objective $B_1 < 1$. Thus the maximum is one.

10.2.3 Equality

If a feasible nongeometric sequence had objective one, the strict-convexity perturbation in the localization argument would produce a feasible sequence with objective greater than one, which is impossible. Every proper geometric truncation has objective strictly less than one by the tail estimate. The only remaining equality case is $L = n$, $t = 1$, and $u_k = 1$ for every k.

Normalization then forces $r_k = 1$ and $P = B$. This proves the equality characterization in the theorem.

The proof applies to Lorentzian inputs, not only to real-stable inputs. Properly truncated geometric likelihood ratios need not give real-rooted binary polynomials. Restriction to Poisson-binomial distributions would therefore exclude some extremal candidates.

10.2.4 Polynomial consequence

Consider a homogeneous polynomial

$$p(x) = \sum_{k=0}^{n} a_k x_1^k (x_2 + \cdots + x_n)^{n-k}. \tag{110}$$

Write $m = n - 1$. Normalization $p(\mathbf{1}) = 1$ gives the marginal probabilities

$$P_k = a_k m^{n-k}.$$

Balance is equivalent to $\mathbb{E}_P K = 1$: the other m coordinates are symmetric, and homogeneity then makes each of their mean exponents one as well.

The explicit coefficient condition for (110) is log-concavity, with interval support, of $a_k/\binom{n}{k}$. Equivalently, $r_k = P_k/B_k$ is log-concave. This is the binary Lorentzian criterion applied to x_1 and the sum of the other variables, as defined above.

The central coefficient and squared Fischer norm satisfy

$$\frac{[x_1 \cdots x_n]p}{c_n} = r_1, \qquad \frac{N(p)}{c_n} = \sum_k B_k r_k^2. \tag{111}$$

For the second identity, the contribution with exponent k on x_1 is

$$k!(n-k)!\, a_k^2 m^{n-k}.$$

Thus (99) proves

$$\big([x_1 \cdots x_n]p\big)^2 \geq c_n N(p)$$

for this entire all-order Lorentzian family. Equality is the uniform power $p = ((x_1 + \cdots + x_n)/n)^n$.

10.3 Symmetric reference laws

10.3.1 Statement

Let $N \geq 1$ and let B be a strictly positive probability mass function on $\{-N, \ldots, N\}$ such that $B_{-j} = B_j$. Let P be a mean-zero probability mass function on the same set such that $r_j = P_j/B_j$ is log-concave with interval support. Then

$$\boxed{\sum_{j=-N}^{N} B_j r_j^2 \leq r_0^2.} \tag{112}$$

Equality holds exactly when $P = B$.

Notice that B itself need not be log-concave. The only shape assumption is log-concavity of the likelihood ratio P/B.

10.3.2 Proof

Set $u = r/r_0$, which is legitimate because interval support and mean zero imply $P_0 > 0$. By the localization lemma, a maximizer is either supported at zero or has the form

$$u_j = e^{sj} \mathbf{1}_{\{-a \leq j \leq b\}}$$

for integers $a, b \geq 1$. A singleton has objective $B_0 < 1$.

If $s = 0$, balance and strict positivity of the symmetric reference force $a = b$. The objective is the reference probability of this interval, at most one, with equality only for the full interval.

Suppose $s > 0$. If $a \leq b$, every symmetric pair in the interval contributes a strictly positive amount to the mean, as do any unpaired positive terms. Therefore balance forces $a > b$. The mean equation becomes

$$\sum_{j=1}^{b} 2jB_j \sinh(sj) = \sum_{j=b+1}^{a} jB_j e^{-sj}. \tag{113}$$

Subtract the reference mass of the interval from the squared-norm objective. Pairing j and $-j$ gives the exact identity

$$\Phi_B(u) - B([-a,b]) = \sum_{j=1}^{b} 4B_j \sinh^2(sj) - \sum_{j=b+1}^{a} B_j(1-e^{-2sj}). \tag{114}$$

The function $x \mapsto \sinh(sx)/x$ is strictly increasing on $x > 0$: its derivative has numerator $sx\cosh(sx) - \sinh(sx)$, whose derivative in sx is positive. Thus

$$\begin{aligned}
\sum_{j=1}^{b} 4B_j \sinh^2(sj) &\le \frac{2\sinh(sb)}{b} \sum_{j=1}^{b} 2jB_j \sinh(sj) \\
&= \frac{2\sinh(sb)}{b} \sum_{j=b+1}^{a} jB_j e^{-sj} \\
&< \sum_{j=b+1}^{a} 2B_j \sinh(sj) e^{-sj} \\
&= \sum_{j=b+1}^{a} B_j(1-e^{-2sj}).
\end{aligned} \tag{115}$$

The strict inequality follows from $j > b$ and the presence of at least one unpaired term. Equations (114) and (115) imply

$$\Phi_B(u) < B([-a,b]) \le 1.$$

For $s < 0$, reflect the indices and apply the same argument. Consequently, the only geometric candidate with objective one is $u \equiv 1$ on the full interval. A nongeometric equality case would admit the strictly improving two-sided perturbation from the localization lemma, contradicting the bound. Therefore equality forces $u \equiv 1$, and probability normalization gives $r \equiv 1$, or $P = B$. This proves (112).

10.3.3 Central binomial consequence

Take $n = 2m$ and $B = \mathrm{Bin}(2m, 1/2)$. It is symmetric about its integer mean m. Thus every mean-m probability mass function P with log-concave likelihood ratio $r = P/B$ satisfies

$$\sum_{k=0}^{2m} B_k r_k^2 \le r_m^2,$$

with equality only for $P = B$. This proves the arbitrary-order central-mean case, in addition to the mean-one case above and its reflection at mean $n-1$.

10.3.4 Exact examples

In these examples, the scalar ratio means $r_m^2 / \sum_k B_k r_k^2$, where m is the mean.

For $B = \mathrm{Bin}(4, 1/2)$, consider support $\{1,2,3\}$. Symmetry gives $t = 1$ and

$$\Phi_B(u) = B_1 + B_2 + B_3 = \frac{7}{8}.$$

After probability normalization, the likelihood ratio is constant $8/7$ on this interval; the scalar ratio is $8/7 > 1$.

For a reference law which is not log-concave, take

$$(B_{-2}, B_{-1}, B_0, B_1, B_2) = \frac{1}{20}(6, 1, 6, 1, 6).$$

It is symmetric and strictly positive, but $B_{-1}^2 < B_{-2}B_0$. The theorem still applies to every mean-zero P with log-concave ratio P/B. For example, restriction to $\{-1, 0, 1\}$ has $t = 1$, objective $2/5$, and scalar ratio $5/2$.

10.3.5 Two-group polynomial consequence

Let $m + \ell = n$, and set

$$X = x_1 + \cdots + x_m, \qquad Y = x_{m+1} + \cdots + x_n, \qquad p(x) = \sum_{k=0}^{n} a_k X^k Y^{n-k}.$$

For a normalized polynomial, define

$$P_k = a_k m^k \ell^{n-k}, \qquad B_k = \binom{n}{k} \left(\frac{m}{n}\right)^k \left(\frac{\ell}{n}\right)^{n-k}.$$

Balance of all n variables is equivalent to $\mathbb{E}_P K = m$. The Lorentzian condition is log-concavity with interval support of $a_k/\binom{n}{k}$, equivalently of $r = P/B$. The identities are

$$\frac{[x_1 \cdots x_n]p}{c_n} = r_m, \qquad \frac{N(p)}{c_n} = \sum_{k=0}^{n} B_k r_k^2. \tag{116}$$

Indeed the central coefficient is $a_m m!\ell!$, and the contribution to the squared Fischer norm from total degree k in the first group is

$$k!(n-k)!a_k^2 m^k \ell^{n-k}.$$

Consequently the symmetric-reference theorem proves the target Fischer-norm inequality for every balanced Lorentzian polynomial with two groups of the same size. It does not require a product of linear forms.

For unequal group sizes, (116) is still exact, but the corresponding integer-mean scalar inequality remains unproved. The proof for symmetric reference laws cannot be applied directly to an asymmetric binomial law by an exponential tilt, because that changes the mean about which the balance condition is imposed.

10.4 Duplication and the Gram permanent

There is a second consequence of the mean-one localization argument, beyond the norm inequality itself.

10.4.1 Scalar bridge theorem

Let $n \geq 2$, let $B = \mathrm{Bin}(n, 1/n)$, and let P be a mean-one probability law whose ratio P/B is log-concave with interval support. Write

$$q = \frac{n-1}{n}, \qquad g = B_1 = q^{n-1}.$$

Then

$$\boxed{P_1^2 + 2P_0P_2 \geq (1+q)g^2 \sum_{k=0}^{n} \frac{P_k^2}{B_k}.} \tag{117}$$

Equality holds exactly for $P = B$.

10.4.2 Proof

Set $u_k = (P_k/B_k)/(P_1/B_1)$ as in the mean-one proof. Since $B_0 = qg$ and $B_2 = g/2$, (117) is equivalent to

$$\Psi(u) := (1+q)\sum_k B_k u_k^2 - qu_0u_2 \leq 1. \tag{118}$$

The same compact feasible set and two-sided multiplicative perturbations from the localization argument apply. It remains to check strict convexity of the new quadratic objective. All coordinates except 0 and 2 are diagonal positive terms. The remaining 2×2 block is positive definite exactly when

$$(1+q)^2 B_0 B_2 > \frac{q^2}{4}, \quad \text{or equivalently} \quad 2(1+q)^2 g^2 > q. \tag{119}$$

The elementary bound $(1-1/n)^{n-1} > e^{-1}$ follows from $\log(1-x) > -x/(1-x)$ for $0 < x < 1$. Also $(1+q)^2/q \geq 4$. Therefore

$$\frac{2(1+q)^2 g^2}{q} > \frac{8}{e^2} > 1,$$

which proves (119). Thus Ψ is strictly convex.

The localization proof now forces a maximizer to be a geometric likelihood ratio on an initial interval $\{0, \ldots, L\}$, $L \geq 2$, or the singleton at 1. In the geometric case, $u_0u_2 = 1$, so

$$\Psi(u) = (1+q)\Phi_B(u) - q \leq 1$$

by the already proved mean-one theorem. Its strictness on every proper geometric truncation gives strictness here as well.

For the singleton, $\Psi(u) = (1+q)g < 1$. Indeed Bernoulli's inequality gives

$$g^{-1} = \left(1 + \frac{1}{n-1}\right)^{n-1} \geq 2,$$

while $q < 1$. Strict convexity and the same perturbation argument rule out any further equality case. Hence equality forces $P = B$.

10.4.3 Polynomial form

For a balanced normalized Lorentzian polynomial of the form

$$p(x) = \sum_{k=0}^{n} a_k x_1^k (x_2 + \cdots + x_n)^{n-k},$$

(117) is exactly the coefficient inequality

$$\boxed{2^n [x_1^2 \cdots x_n^2] p(x)^2 \geq \frac{(2n)!}{n!\, n^n} N(p).} \tag{120}$$

To verify the constants directly, put $m = n - 1$ and $P_k = a_k m^{n-k}$. The left side of (120) is

$$2(2n-2)!(a_1^2 + 2a_0 a_2),$$

while $N(p) = c_n \sum_k P_k^2 / B_k$. Substitution gives exactly (117). Equality in (120) is the uniform power $((x_1 + \cdots + x_n)/n)^n$.

For a matrix product polynomial, the left side of (120) is $\mathrm{per}(A \otimes E_2)$, where E_2 is the all-ones 2×2 matrix. Thus (120) compares the permanent of the duplicated matrix with its Gram permanent for this polynomial family. No extension to arbitrary two-group or multivariate Lorentzian polynomials is asserted here.

10.5 The remaining general integer-mean reduction

The localization statement also gives an exact reduction for any interior integer mean m and $B = \mathrm{Bin}(n, m/n)$. For a law P of mean m with log-concave ratio $r = P/B$ and interval support, the proposed inequality is

$$\sum_k B_k r_k^2 \leq r_m^2.$$

Set $u = r/r_m$. The inequality holds for every such law if and only if

$$\sum_{k=a}^{b} B_k t^{2(k-m)} \leq 1$$

for every interval $a < m < b$, at its unique balancing root

$$\sum_{k=a}^{b} (k-m) B_k t^{k-m} = 0.$$

The singleton candidate has value $B_m < 1$. The equivalence follows because localization identifies every maximizer, and each geometric candidate is feasible. The mean-one theorem, its reflection, and the symmetric-reference theorem prove the cases $m = 1$, $m = n - 1$, and $n = 2m$. The other cases for arbitrary log-concave likelihood ratios remain unproved here. The Poisson-binomial theorem in Section 6 has stronger conclusions on its more restricted class.

11 A two-law scalar reduction in arbitrary order

The rank-two representation also gives an exact scalar form of the strong permanent inequality. A further reduction permits a larger class than Poisson-binomial laws. It reduces that scalar assertion to pairs of geometric likelihood ratios on intervals. The resulting scalar inequality remains unproved. It does not extend the finite-order theorem to every rank-two matrix.

11.1 The permanent and a binomial-noise operator

Fix $n \geq 2$ and $0 < t < n$. Use the representation (75), with $\theta = t/n$. Let P_x, P_y be the corresponding Poisson-binomial laws, and put

$$B(k) = \binom{n}{k}\theta^k(1-\theta)^{n-k}, \qquad r(k) = \frac{P_x(k)}{B(k)}, \qquad s(k) = \frac{P_y(k)}{B(k)}.$$

The reference law B is positive at every index. For functions on $\{0, \dots, n\}$, write $\langle f, g\rangle_B = \sum_k B(k)f(k)g(k)$.

Expand each entry of (75) into its two summands. If k rows use the first summand, the column assignment has $k!(n-k)!$ possible matchings for each selected column set. Summing over row and column sets gives

$$\operatorname{per}(A) = \sum_{k=0}^{n} \frac{k!(n-k)!}{t^k(n-t)^{n-k}} P_x(k)P_y(k).$$

The scalar coefficient is $c_n/B(k)$. Hence

$$\frac{\operatorname{per}(A)}{c_n} = \langle r, s\rangle_B. \tag{121}$$

For $0 \leq \kappa \leq 1$, define a Markov chain on n Bernoulli bits as follows. Each bit is retained with probability κ. Otherwise, replace it by an independent Bernoulli variable with parameter θ. Apply these operations independently across coordinates. The induced transition on the number of successes depends only on the initial number of successes. Write $T_\kappa(k, \ell)$ for this transition probability, and use T_κ also for its operator on functions. The bit transition is reversible for $\operatorname{Bern}(\theta)$. Thus the count transition is reversible for B:

$$B(k)T_\kappa(k, \ell) = B(\ell)T_\kappa(\ell, k).$$

If the initial bit parameters are x_i, their new parameters are $x'_i = \theta + \kappa(x_i - \theta)$. Reversibility shows that their new count likelihood ratio is $T_\kappa r$.

For the law P_y, define

$$\kappa(s) = 1 - \frac{\operatorname{Var}(P_y)}{t(1-t/n)} = \frac{n\sum_j y_j^2 - t^2}{t(n-t)}. \tag{122}$$

This is the same $\kappa_y \in [0, 1]$ as in (78). The Gram matrix is the matrix in (75) made from x and $x' = \theta\mathbf{1} + \kappa(s)(x - \theta\mathbf{1})$. Equation (121) therefore gives

$$\frac{\operatorname{per}(AA^{\mathsf{T}})}{c_n} = \langle r, T_{\kappa(s)}r\rangle_B. \tag{123}$$

Consequently the strong rank-two inequality is exactly

$$\langle r, s\rangle_B^2 \geq \langle r, T_{\kappa(s)}r\rangle_B \tag{124}$$

for Poisson-binomial laws of equal mean. Both identities hold in every order; they do not assert the inequality in every order.

11.2 The larger class and its variance bound

A law P on $\{0,\dots,n\}$ is *ultra-log-concave of order n* if its support is an interval and $P(k)/\binom{n}{k}$ is log-concave there. At fixed θ, this is equivalent to log-concavity of P/B. The two normalizations differ by a positive geometric sequence. Newton's coefficient inequalities for a product of nonnegative binary linear forms show that Poisson-binomial laws belong to this class.

Let $\mathcal{U}_{n,t}$ be the ultra-log-concave laws with mean t. This is a compact set of probability vectors. For example, put $u_k = P(k)/B(k)$. The inequalities

$$u_j^{k-i} \geq u_i^{k-j} u_k^{j-i} \qquad (i<j<k)$$

are closed conditions. They express log-concavity and exclude internal zeros between two positive coefficients. The fixed-mass and fixed-mean conditions are also closed.

Lemma 11.1. *For $P \in \mathcal{U}_{n,t}$ and every convex function φ on the index interval,*

$$\mathbb{E}_P\varphi(K) \leq \mathbb{E}_B\varphi(K).$$

In particular, $0 \leq \mathrm{Var}(P) \leq t(1-t/n)$.

Proof. The positive indices of $P-B$ are the indices where $P/B>1$. Log-concavity makes them an interval. Thus $P-B$ has the sign pattern negative, positive, negative, with zeros permitted. If there are no positive entries, the equal total masses imply $P=B$. Otherwise let a,b be the first and last positive indices. For $a<b$, take the secant line L through $(a,\varphi(a))$ and $(b,\varphi(b))$. Convexity gives $\varphi - L \leq 0$ between those indices and $\varphi - L \geq 0$ outside them. Hence $(P-B)(\varphi-L)\leq 0$ at every index. For $a=b$, use a supporting line to the convex function at a instead. The same product inequality holds. The sum of $(P-B)L$ is zero because P and B have equal mass and mean. Summing proves the expectation comparison. Take $\varphi(k)=(k-t)^2$ for the variance bound. □

For arbitrary $P,Q \in \mathcal{U}_{n,t}$, set $r=P/B$, $s=Q/B$, and

$$\kappa(s) = 1 - \frac{\mathrm{Var}(Q)}{t(1-t/n)}.$$

The lemma makes this a parameter in $[0,1]$. It is affine in Q on the fixed-mean set. The proposed extension is (124) for all such P,Q. This extension is unproved.

11.3 Separate convexity of the square-root deficit

Define

$$\mathcal{D}(r,s) = \sqrt{\langle r, T_{\kappa(s)} r\rangle_B} - \langle r,s\rangle_B.$$

Since the inner product of the nonnegative functions r,s is nonnegative, the target (124) is equivalent to $\mathcal{D}(r,s)\leq 0$.

To obtain the spectrum of T_κ, take independent bits X_i with parameter θ and put $\xi_i = (X_i-\theta)/\sqrt{\theta(1-\theta)}$. The functions

$$K_k = \binom{n}{k}^{-1/2} \sum_{|S|=k} \prod_{i\in S} \xi_i, \qquad 0\leq k\leq n,$$

depend only on the number of successes. They are the normalized Krawtchouk polynomials. Independence makes products from distinct subsets orthogonal, with squared norm one. Thus

$K_0, \dots, K_n$ form an orthonormal basis for the count space with weight B. The conditional mean of an updated centered bit is $\kappa\xi_i$. Independence of the updates therefore gives

$$T_\kappa K_k = \kappa^k K_k.$$

All eigenvalues are nonnegative, so T_κ is positive semidefinite. For fixed s, the first term of $\mathcal{D}$ is a seminorm of r. Subtracting a linear functional preserves convexity.

Mass one and mean t imply $r = 1 + \sum_{k=2}^{n} a_k K_k$. The coefficient of $K_1 = (K - t)/\sqrt{t(1 - t/n)}$ vanishes. Therefore

$$\begin{aligned}\sqrt{\langle r, T_\kappa r\rangle_B} &= \sqrt{1 + \sum_{k=2}^{n} a_k^2 \kappa^k} \\ &= \left\|(1, |a_2|\kappa, |a_3|\kappa^{3/2}, \dots, |a_n|\kappa^{n/2})\right\|_2.\end{aligned}$$

Every coordinate is nonnegative and convex in $\kappa \geq 0$. The Euclidean norm is convex and increases in each coordinate on the nonnegative orthant. The displayed function is therefore convex in κ. Since $\kappa(s)$ is affine, $\mathcal{D}$ is convex in s for fixed r. These arguments apply on the convex hull of $\mathcal{U}_{n,t}$, where the variance bound remains valid by affinity. They prove separate convexity, without a claim of joint convexity.

11.4 Localization to two support intervals

Lemma 11.2 (Fixed-mean ultra-log-concave localization)**.** *Let F be a continuous functional on $\mathcal{U}_{n,t}$ that extends convexly to its convex hull. Then F has a maximizer whose likelihood ratio P/B is geometric on its support interval. A singleton is permitted when t is an integer.*

Proof. Compactness gives a maximizer. Among the maximizers, choose one that maximizes the strictly convex quantity $\sum_k P(k)^2$. Write its positive support as $[a, b]$ and $u = P/B$. At each equality index $u_k^2 = u_{k-1}u_{k+1}$ impose $2h_k = h_{k-1} + h_{k+1}$. If some interior log-concavity inequality is strict, these independent second-difference equations leave a solution space of dimension at least three. The further equations

$$\sum_k P(k)h_k = 0, \qquad \sum_k kP(k)h_k = 0$$

therefore leave a nonzero solution h. For sufficiently small $\varepsilon > 0$, put $P^{\pm}(k) = P(k)(1 \pm \varepsilon h_k)$ on the same support. They preserve positivity, mass, and mean. At an equality index, their log-concavity gap is

$$u_k^2\varepsilon^2(h_k^2 - h_{k-1}h_{k+1}) = u_k^2\varepsilon^2(h_k - h_{k-1})^2 \geq 0.$$

The strict inequalities remain true for small ε. Thus both perturbed laws belong to $\mathcal{U}_{n,t}$ and their average is P. Convexity and maximality force both to maximize F. Strict convexity of $\sum_k P(k)^2$ makes one have a larger auxiliary value, a contradiction. Every interior inequality must therefore be equality. This gives a geometric likelihood ratio. Supports of one or two points already have the required form. □

Apply this lemma first to one argument of a maximizing pair for $\mathcal{D}$, and then to the other. Each replacement preserves the global maximum. The proposed extension is therefore equivalent to the same inequality for pairs

$$P_I(k) = \frac{B(k)e^{\lambda_I k}\mathbf{1}_{\{k\in I\}}}{Z_I}, \qquad Z_I = \sum_{k\in I} B(k)e^{\lambda_I k}, \tag{125}$$

where both laws have mean t. For a nontrivial interval $I = [a, b]$, this requires $a < t < b$. There is then exactly one finite λ_I. Indeed, the derivative of the mean with respect to λ_I is the positive variance, and its limits are a, b. When t is an integer, the singleton at t is also included, with $\lambda_I = 0$. It covers every limiting case in which the mean reaches a support endpoint.

For precision, put $M_\kappa(k, \ell) = B(k)T_\kappa(k, \ell)$. The joint generating polynomial of the initial and updated counts is

$$\sum_{k,\ell} M_\kappa(k, \ell) z^k w^\ell = \big[(1-\theta)^2 + \kappa\theta(1-\theta) + \theta(1-\theta)(1-\kappa)(z+w) + (\theta^2 + \kappa\theta(1-\theta))zw\big]^n. \tag{126}$$

The four terms are the four joint probabilities for one bit and its update. For two interval laws P_I, P_L, set $\kappa_L = 1 - \mathrm{Var}(P_L)/(t(1-t/n))$. After canceling the factor Z_I^{-2}, the remaining inequality is exactly

$$\left[\sum_{k \in I \cap L} B(k) e^{(\lambda_I + \lambda_L)k}\right]^2 \geq Z_L^2 \sum_{k,\ell \in I} M_{\kappa_L}(k, \ell) e^{\lambda_I(k+\ell)}. \tag{127}$$

Inequality (127) **is unproved in general.** There are finitely many interval pairs for each fixed n, t, but t is still continuous. If either law is the full binomial law, the target is equality. This follows from its likelihood ratio one and from $T_0 r = \langle r, 1\rangle_B = 1$. These cases check the normalization; they do not establish the other pairs.

A failure of (127) would refute this larger scalar extension. It would not by itself give a matrix counterexample, since the interval laws need not be Poisson-binomial. The unsquared contraction from Section 8 is not asserted for this larger class. The accompanying exact rational check verifies the permanent identity, variance identity, Gram identity, and count-law update in orders three, four, and five. One of its means is 5/3. Those checks support the implementation of the identities; the proofs above establish the identities in general.

12 Coordinate deletion and exact finite certificates

Coordinate deletion extracts the coefficient of one variable. A deletion inequality holds for matrix-product polynomials of degree three and for two-shift cycles of every order, as proved below. The unrestricted deletion inequality remains unproved.

For a homogeneous nonnegative polynomial q of degree d in d variables, define

$$\mathrm{cap}(q) = \inf_{x_i > 0} \frac{q(x)}{x_1 \cdots x_d}.$$

A positive diagonal matrix D is capacity-balancing when $q(Dx)/\det D$ is balanced. Here balance means that each derivative at $\mathbf{1}$ equals the value at $\mathbf{1}$. The polynomial need not be normalized to value one. At a finite capacity minimum, these are the logarithmic critical-point equations. For a normalized polynomial they agree with the definition in Section 10.

The squared Fischer norm N uses the same definition at each degree and number of variables. If $q_j = [x_j]p$, write

$$H_j = \frac{N(q_j(D_j x))}{(\det D_j)^2}$$

when a capacity-balancing diagonal exists. The degree-three case below has a specified limiting expression when a finite diagonal does not exist. The rank-two closure section treats its own limiting cases. No general minimum-norm interpretation of H_j is assumed.

The cubic deletion inequality is

$$\frac{1}{3}\sum_{j=1}^{3} H_j \geq \frac{4}{9} N(p_A)$$

for every $A \in \Omega_3$, with equality exactly at J_3. Its proof follows.

12.1 The cubic deletion theorem

Write

$$p_A(x) = \prod_{i=1}^{3} (Ax)_i, \qquad b = [x_1 x_2 x_3] p_A = \operatorname{per} A.$$

For each j, the deletion $q_j = [x_j] p_A$ is a binary quadratic

$$q_j(x_k, x_\ell) = a_j x_k^2 + b x_k x_\ell + c_j x_\ell^2.$$

When $a_j, c_j > 0$, choose diagonal scales with product one and with $a_j d_k^2 = c_j d_\ell^2 = \sqrt{a_j c_j}$. The scaled quadratic is balanced and realizes the capacity infimum, so its squared Fischer norm is

$$H_j = 2a_j c_j + b^2 + 2a_j c_j = b^2 + 4a_j c_j.$$

If one of a_j, c_j vanishes, this expression is the limit of the balanced norms. A finite balancing diagonal need not exist in that degenerate case. The final polynomial inequality extends to it by continuity.

The permanent Cauchy–Binet identity gives

$$N(p_A) = \operatorname{per}(AA^T).$$

Since $g(n) = ((n-1)/n)^{n-1}$, and in particular $g(3) = 4/9$, the average-deletion statement $\frac{1}{3}\sum_j H_j \geq g(3) N(p_A)$ is exactly

$$3\sum_j H_j - 4N(p_A) = 9b^2 + 12\sum_j a_j c_j - 4\operatorname{per}(AA^T) \geq 0. \tag{128}$$

12.2 Birkhoff parametrization and chamber coverage

List the six permutation matrices of order three as $P_0, \ldots, P_5$, with $P_0 = I$, and write

$$A = \sum_{r=0}^{5} t_r P_r, \qquad t_r \geq 0.$$

For a doubly stochastic matrix the weights can be chosen with sum one. The polynomial in (128) is homogeneous of degree six in the weights, so the coefficient calculation need not impose this normalization. It proves the inequality for every common row and column sum.

Given an ordering π of the six weights, set

$$t_{\pi(r)} = z_0 + \cdots + z_r, \qquad z_r \geq 0.$$

The 720 chambers cover the nonnegative orthant, including ties. A row permutation of A leaves p_A unchanged because it only reorders its linear factors. Its action on the six permutation matrices is transitive. Therefore whichever Birkhoff vertex has the smallest weight can be sent to P_0. This justifies the primary reduction to the 120 orderings with $\pi(0) = 0$.

Base-seven monomial encoding is exact because every exponent is at most six. Every intermediate product also has total degree at most six, so addition of exponents cannot cause a carry between positions.

12.3 Independent exact verification

The independent checker uses a different construction:

- It builds the matrix directly in gap coordinates, without expanding in raw Birkhoff weights and without applying the primary triangular substitution.
- It computes $N(p_A)$ directly as $\mathrm{per}(AA^T)$, rather than summing factorial-weighted squares of product-polynomial coefficients.
- It obtains a_j, c_j from the three possible choices of the unique row assigned to column j.
- It uses base-eight exponent encoding and checks all 720 chambers without relying on the symmetry reduction.

The program is `deletion_n3_independent_audit.py`. Its recorded output has the corresponding `.json` name. The primary program is `deletion_n3_certificate.py`. The program `deletion_n3_crosscheck.py` gives a second independent construction with exponent tuples. The three implementations agree exactly.

The exact integer computations give the following results:

Quantity	Value
Chambers checked	720
Positive coefficient positions checked	324,360
Smallest positive coefficient	5
Primary chamber counts and minima matched	120 of 120
Independent scalar integer point checks	5,760

The positive coefficient counts per chamber have the following histogram:

Number of positive coefficients	Number of chambers
411	72
452	72
454	144
455	144
456	288

The scalar checks evaluate the generated polynomial at nonnegative integer gap vectors and compare it with direct scalar permanents and the explicit binary-quadratic formula for H_j. They do not reuse the primary coefficient-norm expansion.

The coefficient signs establish the inequality on each whole nonnegative parameter region.

12.4 Equality in the order-three certificate

The independent computation also checks the coefficients of the pure powers z_r^6. In every chamber:

- The coefficient of z_0^6 is zero.
- The coefficients of $z_1^6, z_2^6, z_4^6, z_5^6$ are positive.
- The coefficient of z_3^6 is positive in 648 chambers and zero in 72 chambers.
- In each of those 72 chambers, the three permutation matrices in the tail sum to the all-ones matrix E.

Since all coefficients are nonnegative, equality in (128) forces $z_1 = z_2 = z_4 = z_5 = 0$. If z_3 is nonzero, its coefficient must vanish, and its tail is therefore E. The uniform baseline is $\sum_r P_r = 2E$. In either case, A is a scalar multiple of E.

Conversely, a scalar multiple of E gives equality directly. Thus equality in the average-deletion certificate holds exactly at J_3 among doubly stochastic matrices.

12.5 Conditional induction: normalization and closure

For a homogeneous degree-n polynomial p with nonnegative coefficients, call it balanced when

$$\partial_i p(\mathbf{1}) = p(\mathbf{1}) \quad \text{for every } i.$$

The coefficient distribution then has mean exponent one in each coordinate. Weighted AM–GM gives $\operatorname{cap}(p) = p(\mathbf{1})$.

Suppose the following lemma were proved for every $n \geq 2$ and every balanced real-stable polynomial with strictly positive degree-n coefficients:

$$\max_j H_j \geq g(n)N(p), \qquad g(n) = \left(\frac{n-1}{n}\right)^{n-1}. \tag{129}$$

For an index j supplied by (129), put

$$r(x) = \frac{q_j(D_j x)}{\det D_j}, \qquad q_j = [x_j]p.$$

Then:

1. Differentiation followed by specialization $x_j = 0$ preserves real stability, or gives zero. Here strict coefficient positivity prevents zero.
2. q_j has nonnegative coefficients, degree $n-1$, and $n-1$ variables.
3. Positive diagonal scaling and positive scalar multiplication preserve stability and coefficient positivity.
4. The capacity critical-point equations say that r is balanced.
5. The central coefficient of r equals the central coefficient of p: diagonal scaling multiplies that coefficient by exactly $\det D_j$.
6. $N(r) = H_j$ by definition.

Therefore the inductive inequality in degree $n-1$ would give

$$\big([x_1 \cdots x_n]p\big)^2 \;\geq c_{n-1}H_j \;\geq c_{n-1}g(n)N(p) \;= c_nN(p).$$

No normalization to $p(\mathbf{1}) = 1$ is necessary, since all these inequalities are homogeneous of degree two in the polynomial coefficients. If the lemma is instead stated with that normalization, divide by $p(\mathbf{1})$ and restore the scalar afterward.

The degree-one base case is equality. In degree two, a balanced binary quadratic has the form $ax^2 + bxy + ay^2$. Real stability gives $b^2 \geq 4a^2$, and hence

$$b^2 \geq \frac{1}{2}(b^2 + 4a^2) = c_2N(p).$$

This also shows that the degree-three matrix-product certificate, together with the valid binary base case, independently implies the strong permanent inequality for all doubly stochastic matrices of order three.

An induction also requires closure of the polynomial class under deletion. The polynomial q_j is generally a sum of products, not a single product of nonnegative linear forms. A theorem about matrix products in degree three does not establish (129) for all stable polynomials in degree three, and therefore does not by itself enable the next induction step in degree four.

12.6 Boundary cases can be handled without assuming finite balancing

For a nonzero homogeneous polynomial of degree d in d variables with strictly positive coefficients, capacity balancing exists. On the logarithmic gauge $\sum_i y_i = 0$, the function

$$\log q(e^y) - \sum_i y_i$$

is coercive because q contains every pure power with a positive coefficient. It therefore attains a minimum. The Hessian is a covariance matrix; with full monomial support it is positive definite on the gauge subspace. The balancing diagonal is unique up to common scalar multiplication, which does not change H_j.

For a general nonnegative balanced real-stable polynomial, use the approximation

$$p_\varepsilon(x) = p(M_\varepsilon x), \qquad M_\varepsilon = (1-\varepsilon)I + \varepsilon J_n, \quad 0 < \varepsilon < 1.$$

This approximation has all the required properties:

- M_ε is strictly positive and doubly stochastic.
- It maps vectors with all imaginary parts positive to vectors with all imaginary parts positive, so p_ε is real stable.
- The chain rule and double stochasticity preserve balance.
- It preserves degree and makes every degree-n coefficient strictly positive.
- $p_\varepsilon \to p$ coefficientwise as $\varepsilon \downarrow 0$.

Thus one can prove the conditional induction for strictly positive coefficient polynomials and pass directly to the final inequality by continuity of its central coefficient and squared Fischer norm. This avoids asserting that a finite capacity-balancing diagonal exists for every degenerate deletion polynomial. It also avoids the unjustified perturbation $p + \varepsilon(\sum x_i)^n$, since arbitrary sums of stable polynomials need not be stable.

12.7 The permanent-gradient identity

For a matrix product $p_A = \prod_i (Ax)_i$, the following identity computes H_j without storing the full coefficient array. It holds for every specified positive diagonal D_j, independently of whether it balances capacity. Fix j, write

$$a = A[:, j], \qquad B = A[:, \{0, \dots, n-1\} \setminus \{j\}]D_j, \qquad G = BB^T.$$

Differentiating the product gives

$$q_j(D_j x) = \sum_i a_i \prod_{r \neq i} (Bx)_r.$$

The Fischer inner product of the products indexed by i, k is the permanent of G with row i and column k deleted. Consequently,

$$H_j = \frac{a^T (\nabla \operatorname{per}(G)) a}{(\det D_j)^2}. \tag{130}$$

Here the gradient is the entrywise matrix derivative. The deletion polynomial can also be written using $L_i(x) = \sum_{k \neq j} a_{ik} x_k$,

$$q_j(x) = \left(\prod_i L_i(x) \right) \left(\sum_i \frac{a_{ij}}{L_i(x)} \right). \tag{131}$$

The rational expression (131) is first read where all $L_i(x)$ are nonzero. Multiplication by their product gives the polynomial identity, which extends to every x. The permanent-gradient formula follows from the same coefficient pairing as Lemma 2.1, now with $n-1$ row forms. It is an identity and does not supply a lower bound for H_j.

12.8 Deletion on a two-shift cycle

This section proves $H_j \geq g(n) N(p_n)$, where $g(n) = ((n-1)/n)^{n-1}$, for every two-shift cycle in every order $n \geq 2$. The family already satisfies the strong permanent inequality. This additional result gives its exact deletion behavior.

Take

$$p_n(x) = \prod_{i \bmod n} (ax_i + bx_{i+1}), \qquad a, b \geq 0.$$

Deleting one variable joins its adjacent factors:

$$q_j = (a^2 x_{j-1} + b^2 x_{j+1}) \prod_{i \notin \{j-1, j\}} (ax_i + bx_{i+1}). \tag{132}$$

For $a, b > 0$, diagonal balancing turns (132), after division by the product of diagonal entries, into the length-$n-1$ cycle product with weights

$$a' = a^{n/(n-1)}, \qquad b' = b^{n/(n-1)}. \tag{133}$$

To verify this, view (132) as a cycle of $n-1$ two-term forms with coefficient pairs α_i, β_i. Their products are a^n, b^n. Choose variable scales so that every transformed ratio of coefficients is a'/b'; this is possible because the product of the required ratios is one. The product of row multipliers is exactly the product of variable scales. The resulting circulant product is balanced, and therefore realizes the capacity minimum.

Let $N_n(a,b) = N(p_n)$. With $d = a^2 + b^2$, let λ, μ be the roots of $z^2 - dz - a^2b^2 = 0$. The cyclic tridiagonal permanent formula gives

$$N_n(a,b) = 2(ab)^n + \lambda^n + \mu^n. \tag{134}$$

At $n = 1$, this reads $2ab + \lambda + \mu = (a+b)^2$, which also agrees with the coincident-shift one-dimensional matrix.

Thus (133) gives the exact identity

$$H_j = N_{n-1}\big(a^{n/(n-1)}, b^{n/(n-1)}\big). \tag{135}$$

Formula (134) gives

$$H_j \geq (\sqrt{2} - 1) N_n(a,b) \quad (n \geq 2). \tag{136}$$

By homogeneity assume $ab = 1$ and put $t = \log(a/b)$. Define

$$\phi(t) = \operatorname{arsinh}(\cosh t), \qquad h = \phi(0) = \log(1 + \sqrt{2}).$$

Then

$$\phi''(t) = \frac{2\cosh t}{(1 + \cosh^2 t)^{3/2}} > 0,$$

so with $u = n\phi(t)$ and $v = (n-1)\phi(nt/(n-1))$, convexity yields $v \geq u - h$. Write $r = e^{-h} = \sqrt{2} - 1$. Formula (134) becomes $N_n = 2 + e^u + (-1)^n e^{-u}$, and its analog for (135) has exponent v and the opposite parity. These expressions increase with their nonnegative exponent. If n is odd,

$$\big(2 + e^{u-h} + e^{-u+h}\big) - r\big(2 + e^u - e^{-u}\big) = 2(1-r) + e^{-u}(r + r^{-1}) > 0.$$

If n is even, the corresponding difference is

$$2(1-r) - e^{-u}(r + r^{-1}) = (r + r^{-1})(r - e^{-u}) \geq 0,$$

because $u \geq h$. This proves (136), with zero weights handled by continuity.

Since $g(n)$ decreases with n, and $g(5) = 256/625 < \sqrt{2} - 1$, (136) proves $H_j \geq g(n) N(p_n)$ for this family in every order $n \geq 5$. The small orders follow directly:

- For $n = 2$, $H = (a^2 + b^2)^2 \geq \frac{1}{2}((a^2 + b^2)^2 + 4a^2b^2) = g(2)N_2$.
- For $n = 3$, after division by a^3b^3, $9H - 4N_3$ equals $5z^3 - 39z + 46$, where $z = a/b + b/a \geq 2$. Its derivative is positive for $z \geq 2$, and its value at two is eight.

- For $n = 4$, with $ab = 1$, set $x = \cosh(2t/3) = 1 + y$, $y \geq 0$. The difference $64H - 27N_4$ is

$$52 + 1848y + 11892y^2 + 31424y^3 + 31968y^4 + 14208y^5 + 2368y^6,$$

  which is positive.

Accordingly the capacity-balanced deletion inequality is proved on every two-shift cycle. No analogous closure under deletion has been established for three or more arbitrary shifts: the deletion becomes a sum of products, and does not generally remain a single circulant product.

13 Rank-two deletion and balanced limits

The finite theorem in Section 8 does not give a proof in all orders. Two closure properties are relevant to an induction on the number of variables. First, extraction of one copy of a variable preserves the rank-two product class. This remains true after capacity balancing, even when no finite balancing scale exists. Second, a column with a zero entry gives a real-rooted collision polynomial. Neither result proves the deletion norm inequality needed for that induction.

Write $N(p) = \langle p, p \rangle_F$ for the squared Fischer norm. The same coefficient formula applies to any degree and any number of variables. The proof of Lemma 2.1 also permits rectangular factor matrices: the number of linear factors determines the order of the Gram permanent.

13.1 Extraction of one copy of a variable

For a polynomial p, the notation $[x_j]p$ means the coefficient of exactly one copy of x_j. The result is a polynomial in the remaining variables. It is obtained by differentiating in x_j and then setting $x_j = 0$.

Lemma 13.1 (Binary directional derivative)**.** *Let $F(u, v) = \prod_{i=1}^{n}(c_i u + d_i v)$, where $c_i, d_i \geq 0$, and let $a, b \geq 0$. Then $(a\partial_u + b\partial_v)F$ is either zero or a product of $n - 1$ nonnegative binary linear forms, with a positive scalar factor.*

Proof. The assertion is immediate if $n = 1$. First assume $n \geq 2$, all c_i, d_i are positive, their ratios are distinct, and $(a, b) \neq (0, 0)$. Order the roots of $F(t, 1)$ as $r_1 < \cdots < r_n < 0$. For

$$G(t, 1) = \sum_{i=1}^{n}(ac_i + bd_i) \prod_{k \neq i}(c_k t + d_k),$$

the sign at r_i is $(-1)^{n-i}$. Only the summand that omits the factor with root r_i is nonzero there. The signs alternate, so $G(t, 1)$ has a root in each interval (r_i, r_{i+1}). Its degree is $n - 1$ and its leading coefficient is positive. It therefore factors into $n - 1$ linear factors with negative roots. Homogenization gives the asserted nonnegative binary factors.

For zero coefficients or repeated ratios, use a sequence of positive coefficients with distinct ratios. The derivative polynomials converge coefficientwise to G. Normalize each of their binary factors so its two coefficients sum to one. The normalized factors belong to the compact segment $\{(s, 1 - s) : 0 \leq s \leq 1\}$. The remaining scalar is the value of the derivative at $(1, 1)$. If G is nonzero, this scalar tends to the positive number $G(1, 1)$. A convergent subsequence of the factor lists gives the required factorization of G. This argument includes factors u and v. If the derivative vanishes, it gives the stated zero case. □

Proposition 13.2 (Rank-two deletion closure)**.** *Let $A \geq 0$ be an $n \times n$ matrix of rank at most two. Then $[x_j]p_A$ is either zero or a positive scalar multiple of a product of $n-1$ nonnegative linear forms whose coefficient matrix has rank at most two. If $A \in \Omega_n$, this polynomial is nonzero, and its central coefficient is* $\operatorname{per}(A) > 0$.

Proof. A zero row makes $p_A = 0$, so suppose all rows are nonzero. Positive row normalization changes p_A only by a positive scalar. It also preserves rank. Thus it suffices first to assume that A is row stochastic. Its rows belong to a linear space of dimension at most two. The coordinate-sum functional equals one on every row. Consequently their affine span has dimension at most one. If they are not all equal, let v_1, v_2 be the endpoint rows of their convex hull. There are $\theta_i \in [0,1]$ such that

$$A_{i\bullet} = \theta_i v_1 + (1-\theta_i)v_2.$$

Set $L_1(x) = v_1 \cdot x$, $L_2(x) = v_2 \cdot x$, and $F(u,v) = \prod_i (\theta_i u + (1-\theta_i)v)$. Then $p_A = F(L_1, L_2)$. If all rows are equal, use $L_1 = L_2$ instead.

Write $L_1 = ax_j + L_1'$ and $L_2 = bx_j + L_2'$, where $a, b \geq 0$ and the primed forms omit x_j. The chain rule gives

$$q_j := [x_j]p_A = (a\partial_u + b\partial_v)F(L_1', L_2'). \tag{137}$$

Apply Lemma 13.1. Each resulting linear factor has a nonnegative coefficient vector in the span of the coefficient vectors of L_1', L_2'. The factor matrix therefore has rank at most two. A factor can vanish after substitution; this gives precisely the allowed zero case. For $n \geq 2$, the positive scalar from row normalization can be put into one factor. For $n = 1$, the deletion is a constant, which is covered by the scalar multiple of the empty product.

Now suppose A is doubly stochastic. For any set S of rows, let $N(S)$ be the columns with a positive entry in one of those rows. The total row mass of S is $|S|$. The column sums give $|S| \leq |N(S)|$. Hall's matching theorem gives a permutation that uses only positive entries. Thus $\operatorname{per}(A) > 0$. Finally,

$$\left[\prod_{k \neq j} x_k\right] q_j = [x_1 \cdots x_n]p_A = \operatorname{per}(A) > 0. \tag{138}$$

This excludes the zero polynomial. □

13.2 Capacity and a unique balanced limit

For a nonnegative homogeneous polynomial q of degree m in m variables, define its capacity by

$$\operatorname{cap}(q) = \inf_{x_1,\ldots,x_m > 0} \frac{q(x_1,\ldots,x_m)}{x_1 \cdots x_m}.$$

If its central coefficient is $d > 0$, then

$$0 < d \leq \operatorname{cap}(q) \leq q(\mathbf{1}) < \infty. \tag{139}$$

The first inequality uses the central monomial alone.

Theorem 13.3 (Unique capacity-balanced limit)**.** *Let $m \geq 1$, and let q be a nonnegative homogeneous polynomial of degree m in m variables, with capacity $\gamma > 0$. For $z \in \mathbb{R}^m$ with $\sum_i z_i = 0$, put*

$$q_z(x) = q(e^{z_1}x_1, \ldots, e^{z_m}x_m), \qquad C(z) = q_z(\mathbf{1}).$$

Every sequence with $C(z) \to \gamma$ has the same coefficientwise limit $\bar{q}$. This limit satisfies

$$\bar{q}(\mathbf{1}) = \gamma, \qquad \partial_i \bar{q}(\mathbf{1}) = \gamma \quad (1 \le i \le m), \qquad \bar{q}_{\mathbf{1}} = q_{\mathbf{1}}. \tag{140}$$

In particular, $\bar{q}/\gamma$ is normalized and balanced. The quantity

$$H(q) := N(\bar{q}) \tag{141}$$

is well-defined and equals the limit of $N(q_z)$ along every capacity-minimizing sequence.

Proof. Homogeneity permits determinant-one variable scales, so $\inf_{\sum z_i = 0} C(z) = \gamma$. Write $q_\alpha(z) = q_\alpha e^{\alpha \cdot z}$ for the coefficient of x^α in q_z. For two such scale vectors z, w,

$$\begin{aligned} \sum_\alpha \left(\sqrt{q_\alpha(z)} - \sqrt{q_\alpha(w)}\right)^2 &= C(z) + C(w) - 2C((z+w)/2) \\ &\le C(z) + C(w) - 2\gamma. \end{aligned} \tag{142}$$

Along any capacity-minimizing sequence, the vector of square roots of the coefficients is Cauchy. There are finitely many coefficients, so this vector has a limit. The same inequality, with one point from each of two minimizing sequences, shows that their limits agree. Squaring the coordinates gives a unique coefficientwise limit $\bar{q}$. Its coefficient sum is γ. The central coefficient is unchanged because its scaling multiplier is $e^{\sum_i z_i} = 1$.

For any fixed $h \in \mathbb{R}^m$, the capacity bound gives

$$q_z(e^{h_1}, \dots, e^{h_m}) e^{-\sum_i h_i} \ge \gamma.$$

Pass to the coefficientwise limit. The smooth function $\bar{q}(e^h) e^{-\sum_i h_i}$ has minimum γ at $h = 0$. Its derivative in coordinate i at zero is $\partial_i \bar{q}(\mathbf{1}) - \bar{q}(\mathbf{1})$. This proves the balance identities. The squared Fischer norm is continuous in the finite coefficient vector. This proves the assertion about $H(q)$. □

A finite minimizing scale need not exist. For example, $q(x, y) = x(x + y)$ has capacity one, because $q(x, y)/(xy) = 1 + x/y$. The determinant-one scales (e^{-t}, e^t) give

$$q_t(x, y) = e^{-2t} x^2 + xy \longrightarrow xy.$$

The theorem gives this limit without an assumption of finite attainment. This example concerns the general capacity statement; it is not asserted to be a deletion of a doubly stochastic product.

Corollary 13.4 (Rank-two closure through balancing)**.** *Suppose q is a product of m nonnegative linear forms, with coefficient matrix of rank at most two and $\operatorname{cap}(q) = \gamma > 0$. Then the limit in Theorem 13.3 has the form*

$$\bar{q} = \gamma p_M$$

for some $M \in \Omega_m$ of rank at most two. In particular, every singleton deletion of a doubly stochastic rank-two product has this property.

Proof. Let B be a nonnegative coefficient matrix for the factors of q. Put any positive overall scalar into one row. No row is zero, because $q \neq 0$. For a positive diagonal variable scale D_z, normalize each row of BD_z to sum one. Call the resulting row-stochastic matrix M_z. Its rank is at most two, and the product of the removed row sums is $C(z)$. Hence

$$p_{M_z} = q_z / C(z).$$

Along a capacity-minimizing sequence, take a convergent subsequence of M_z. The row-stochastic matrices form a compact set. Its limit M is nonnegative and row stochastic. It has rank at most two, since this condition is closed: all 3×3 minors vanish. For $m < 3$, the rank condition is automatic. Continuity of the product polynomial gives $p_M = \bar{q}/\gamma$.

For a row-stochastic matrix, $\partial_j p_M(\mathbf{1}) = \sum_i M_{ij}$. The balance identities in Theorem 13.3 therefore make every column sum one. Thus $M \in \Omega_m$. The last assertion follows from Proposition 13.2 and (139). The factor matrix need not be unique. The polynomial limit and its squared Fischer norm are unique, which is all that this conclusion requires. □

13.3 The exact missing step for induction

The row-pair variation in Theorem 9.7 preserves the row space. It therefore also gives a boundary reduction within the rank-at-most-two class. That class is compact inside Ω_n. The closure results above would then permit induction if the following assertion were proved: for every rank-at-most-two matrix $A \in \Omega_n$ on the boundary, some column j with a zero entry satisfies

$$H([x_j]p_A) \geq g_n N(p_A), \qquad g_n = \left(\frac{n-1}{n}\right)^{n-1}. \tag{143}$$

Assertion (143) **is unproved.** It is not a conclusion of the closure or capacity arguments.

To check the proposed implication, put $q_j = [x_j]p_A$ and $\bar{q}_j = \gamma p_M$ as above. The strong inequality at order $n-1$, if available, gives

$$\operatorname{per}(A)^2 = \gamma^2 \operatorname{per}(M)^2 \geq c_{n-1} N(\bar{q}_j) = c_{n-1} H(q_j).$$

Here the central coefficient is preserved, and the factor γ^2 cancels in the norm. Assertion (143) would give

$$\operatorname{per}(A)^2 \geq c_{n-1} g_n N(p_A) = c_n N(p_A).$$

The identity $c_{n-1} g_n = c_n$ is exact. Thus this missing assertion would complete an induction, with the trivial order-one case as its base.

The version that requires the same bound for every column is false. For example, let

$$A = \frac{1}{1500}\begin{pmatrix} 482 & 725 & 293 \\ 478 & 775 & 247 \\ 540 & 0 & 960 \end{pmatrix}.$$

This matrix is doubly stochastic of rank two. Its only zero is in column two. For deletion of column one,

$$q = [x_1]p_A = ay^2 + byz + cz^2, \quad a = \frac{899}{10000}, \quad b = \frac{303539}{1125000}, \quad c = \frac{4797067}{56250000}.$$

For a determinant-one scale (e^t, e^{-t}), the two outer coefficients become ae^{2t} and ce^{-2t}. The minimum coefficient sum occurs when both equal $\sqrt{ac}$. Thus $H(q) = b^2 + 4ac$. Direct rational arithmetic gives

$$g_3 N(p_A) - H(q) = \frac{4550848849393}{1186523437500000} > 0.$$

The failing column has no zero, so this example does not disprove (143).

13.4 A real-rooted collision marginal

Theorem 13.5. *Let $A \geq 0$ have rank at most two and no zero rows. Suppose a nonzero column $a = a_j$ contains a zero entry. Then*

$$h_j(z) = \operatorname{per}(AA^{\mathsf{T}} + (z-1)aa^{\mathsf{T}}) \tag{144}$$

has nonnegative coefficients and only nonpositive real roots. A positive constant polynomial is permitted. If $q_k = [x_j^k]p_A$, then

$$h_j(z) = \sum_{k=0}^{n} k!N(q_k)z^k. \tag{145}$$

The same conclusion holds after arbitrary positive scaling of the other columns.

Proof. For $z \geq 0$, scale column j of A by $\sqrt{z}$. The resulting Gram matrix is $AA^{\mathsf{T}} + (z-1)aa^{\mathsf{T}}$. The coefficient identity in Lemma 2.1, grouped by the exponent of x_j, gives (145). Both sides are polynomials in z, so the identity holds for all real z. It proves coefficient nonnegativity.

A nonnegative rank-one matrix with no zero rows has every nonzero column strictly positive. The stated hypotheses therefore imply rank two. Write $G = AA^{\mathsf{T}} = UU^{\mathsf{T}}$, with U of full column rank two. The vector a belongs to the column space of U, so $a = Uw$ for some $w \neq 0$. Since

$$G - aa^{\mathsf{T}} = A_{-j}A_{-j}^{\mathsf{T}} \succeq 0,$$

full column rank implies $I - ww^{\mathsf{T}} \succeq 0$. Indeed, U^{T} maps onto $\mathbb{R}^2$, so its quadratic form tests every vector in $\mathbb{R}^2$. Consequently

$$0 < s := \|w\|^2 \leq 1. \tag{146}$$

Choose a row i with $a_i = 0$. Its feature vector u_i is nonzero, because $G_{ii} > 0$. It is orthogonal to w. Thus $e_1 = u_i/\|u_i\|$ and $e_2 = w/\|w\|$ form an orthonormal basis of $\mathbb{R}^2$. Let (α_k, β_k) be the coordinates of row u_k in this basis. Both coordinates are nonnegative:

$$\alpha_k = \frac{G_{ki}}{\|u_i\|} \geq 0, \qquad \beta_k = \frac{a_k}{\|w\|} \geq 0.$$

The first inequality uses entrywise nonnegativity of AA^{T}. In these coordinates,

$$G + (z-1)aa^{\mathsf{T}} = \alpha\alpha^{\mathsf{T}} + (1 - s + sz)\beta\beta^{\mathsf{T}}. \tag{147}$$

Define $f(t) = \prod_{k=1}^{n}(\alpha_k + \beta_k t) = \sum_{k=0}^{n} f_k t^k$. No factor is zero, because no row of U is zero. All roots of f are nonpositive real numbers; zero roots and a missing top degree are allowed. Apply the factorial coefficient identity to this product of binary forms. It gives

$$h_j(z) = \sum_{k=0}^{n} k!(n-k)!f_k^2(1 - s + sz)^k = n!R(1 - s + sz), \qquad R(t) = \sum_{k=0}^{n} \frac{f_k^2}{\binom{n}{k}}t^k. \tag{148}$$

The Schur–Szegő composition theorem states that negative-rooted degree-n polynomials are closed under binomially normalized coefficientwise composition. Its application to f with itself shows that R has only nonpositive real roots [13, Proposition 1.5]. For zero coefficients or missing degree, first replace the coefficients of each linear factor by positive values and then take a limit. This gives negative-rooted degree-n approximants. The limit is nonzero because $R(1) > 0$. Thus the same root conclusion holds for the limit, with a lower degree permitted.

By (146), the affine substitution $1 - s + sz$ has positive slope and a nonnegative constant term. A root $r \leq 0$ of R gives a root $(r - 1 + s)/s \leq 0$ of h_j. This proves real-rootedness. Also, $h_j(1) = \operatorname{per}(G) > 0$ because the identity permutation contributes the positive product of the diagonal entries of G. Positive scaling of other columns preserves all the hypotheses, so the proof applies after such a scale. □

The normalized coefficients of h_j form a probability law for the occupancy of column j under the factorial coefficient weights. Real-rootedness makes this a sum of independent Bernoulli variables, with deterministic factors permitted. Its mean need not equal one, even when A is doubly stochastic. For example,

$$A = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/2 & 1/2 \end{pmatrix} \quad \Longrightarrow \quad h_1(z) = \frac{3 + 8z + 4z^2}{64}.$$

Its roots are $-1/2, -3/2$. The corresponding Bernoulli parameters are $2/3$ and $2/5$, so

$$\frac{h_1'(1)}{h_1(1)} = \frac{16}{15} \neq 1.$$

The integer-mean likelihood theorem therefore cannot be applied with target mean one. This real-rootedness result does not prove (143). A zero row would make h_j identically zero, so no normalized law would exist. A zero column gives a constant polynomial. These cases explain the explicit nonvanishing hypotheses in the theorem.

14 Local polynomial inequalities and boundary reductions

The matrix product polynomial places the strong Gram inequality inside a larger polynomial problem. This section proves that problem near the uniform power. It also gives exact reductions and counterexamples that identify the limits of the argument. The neighborhood may depend on the degree. No unrestricted polynomial inequality is asserted.

14.1 The polynomial classes and the deletion functional

A real polynomial is *real stable* if it has no zero when every variable has positive imaginary part. A nonzero homogeneous polynomial is *hyperbolic in direction* e if $p(e) \neq 0$ and $t \mapsto p(x+te)$ has only real zeros for every real x. For a homogeneous polynomial with nonnegative coefficients, hyperbolicity in a strictly positive direction implies real stability. This follows from the hyperbolicity-cone theorem and the stability characterization in Borcea et al. [2, Proposition 4.4 and Theorem 4.5].

For a set S of nonnegative integer vectors with a common coordinate sum, *M-convexity* means the following exchange property. If $\alpha, \beta \in S$ and $\alpha_i > \beta_i$, there is a j with $\alpha_j < \beta_j$ such that both $\alpha - e_i + e_j$ and $\beta + e_i - e_j$ belong to S. The following conditions characterize a *Lorentzian* polynomial. A homogeneous polynomial with nonnegative coefficients is Lorentzian if its support is M-convex and every derivative of order two less than its degree has a Hessian with at most one positive eigenvalue. Degrees zero and one have the usual nonnegative-coefficient convention. This characterization agrees with the closed class defined by Brändén and Huh. Real-stable homogeneous polynomials with nonnegative coefficients belong to this class. The class is closed under limits, differentiation, specialization to zero, and nonnegative linear substitutions [3, Theorems 2.10 and 2.25].

Fix $n \geq 2$ and put

$$m = n - 1, \qquad R(x) = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad p_0 = R^n, \qquad g_n = \left(\frac{n-1}{n}\right)^{n-1}.$$

Thus $c_n = g_n c_m$. Unless stated otherwise, p has degree n in n variables and satisfies

$$p(\mathbf{1}) = 1, \qquad \partial_i p(\mathbf{1}) = 1 \quad (1 \leq i \leq n). \tag{149}$$

These conditions are called normalization and balance. For an unnormalized polynomial, balance means $\partial_i p(\mathbf{1}) = p(\mathbf{1})$. Write

$$a(p) = [x_1 \cdots x_n]p, \qquad N(p) = \langle p, p\rangle_F.$$

The polynomial form of the strong inequality is

$$a(p)^2 \geq c_n N(p). \tag{150}$$

For $p = p_A$, this is $\operatorname{per}(A)^2 \geq c_n \operatorname{per}(AA^{\mathsf{T}})$ by Lemma 2.1.

For a degree-m polynomial q in m variables, define its *capacity* by

$$\operatorname{Cap}(q) = \inf_{x_i > 0} \frac{q(x)}{\prod_i x_i}.$$

A positive diagonal matrix D *balances* q if $q(Dx)$ is balanced. For the deletion $q_j = [x_j]p$, coefficient extraction leaves m variables. Whenever a balancing matrix D_j exists, set

$$H_j(p) = \frac{N(q_j(D_j x))}{(\det D_j)^2}, \quad \mathcal{D}_{\mathrm{av}}(p) = \frac{\sum_j H_j(p)}{n g_n N(p)},$$
$$\mathcal{D}_{\max}(p) = \frac{\max_j H_j(p)}{g_n N(p)}, \qquad \mathcal{S}(p) = \frac{a(p)^2}{c_n N(p)}.$$

A common scalar factor in D_j cancels in H_j. All three ratios equal one at p_0.

Lemma 14.1. *Near p_0, the balancing matrices exist and are unique up to a common scalar factor. In the gauge $\det D_j = 1$, their logarithms and the functions H_j are real analytic in the coefficients. Also, every normalized p satisfies*

$$N(p) = c_n + N(p - p_0). \tag{151}$$

Proof. The uniform deletion is

$$q_{0,j} = g_n \left(\frac{\sum_{i \neq j} x_i}{m}\right)^m.$$

Write $D = \operatorname{diag}(e^{y_i})$ with $\sum_i y_i = 0$. The logarithmic capacity objective is $\Psi_q(y) = \log q(e^y) - \sum_i y_i$. Its gradient is the mean exponent minus $\mathbf{1}$. Its Hessian is the covariance matrix of the coefficient distribution tilted by $e^{\alpha \cdot y}$. For $q_{0,j}$ this distribution is multinomial with m trials and equal cell probabilities. The Hessian at zero is $I - \mathbf{1}\mathbf{1}^{\mathsf{T}}/m$. It is the identity on the gauge subspace $\mathbf{1}^{\perp}$. The implicit-function theorem gives an analytic critical point for nearby coefficients. These coefficients are strictly positive. The pure monomials then make Ψ_q coercive on the gauge subspace, and its Hessian is positive definite there. Thus this critical point is the unique global minimum. For $m = 1$, the gauge subspace is zero and the assertion is immediate. Finally, the coefficients of p_0 give $\langle p_0, h\rangle_F = c_n h(\mathbf{1})$. Apply this identity to $h = p - p_0$ to obtain (151). □

Define the dimensionless distance

$$d_F(p,p_0)=\sqrt{N(p-p_0)/c_n}.$$

The next calculation does not require stability.

Lemma 14.2. *Suppose $p_t=p_0+th+t^2k+o(t^2)$ is a normalized balanced coefficient path. For each j, put*

$$s_j=[x_j]h,\quad a_j=s_j(\mathbf{1}),\quad v_j=\nabla s_j(\mathbf{1})-a_j\mathbf{1},\quad b_j=([x_j]k)(\mathbf{1}).$$

Then $\mathbf{1}^{\mathsf T}v_j=0$, and in the zero-sum logarithmic gauge,

$$\log D_j(t)=-tv_j/g_n+O(t^2), \tag{152}$$

$$H_j(p_t)=g_nc_n+2c_na_jt+\bigl(N(s_j)-2c_m\|v_j\|^2+2c_nb_j\bigr)t^2+o(t^2), \tag{153}$$

$$N(p_t)=c_n+N(h)t^2+o(t^2). \tag{154}$$

Proof. Euler's identity gives $\sum_i\partial_i s_j(\mathbf{1})=ma_j$. Differentiating the capacity equation in Lemma 14.1 gives (152). To compute the second coefficient, write $q_0=q_{0,j}$, $s=s_j$, $v=v_j$, and $d=-v/g_n$. Let $L_d=\sum_i d_ix_i\partial_i$. The first coefficient of the scaled deletion is $s+L_dq_0$. The Fischer identities

$$\langle q_0,f\rangle_F=c_nf(\mathbf{1}),\quad \langle s,L_dq_0\rangle_F=g_nc_m\,d\cdot v,\quad N(L_dq_0)=g_n^2c_m\|d\|^2$$

follow by expanding the uniform power. They imply $N(s+L_dq_0)=N(s)-c_m\|v\|^2$. The second coefficient of the scaled polynomial contains $[x_j]k+L_ds+\frac12L_d^2q_0+L_eq_0$, where $\sum_i e_i=0$ is the second logarithmic displacement. Its inner product with $2q_0$ is

$$2c_nb_j+2c_nd\cdot v+c_ng_n\|d\|^2=2c_nb_j-c_m\|v\|^2.$$

The L_e term vanishes because $\sum_i e_i=0$. This proves (153). Equation (154) follows from (151). □

For a straight path, the quadratic coefficient of $n^{-1}\sum_j H_j-g_nN(p)$ is

$$\mathcal{Q}(h)=\sum_{|\alpha|=n}\left(\frac{k_1(\alpha)}{n}-g_n\right)\alpha!h_\alpha^2-\frac{2c_m}{n}\sum_j\|v_j\|^2, \tag{155}$$

where $k_1(\alpha)=\#\{j:\alpha_j=1\}$. Indeed, $\sum_j N([x_j]h)=\sum_\alpha k_1(\alpha)\alpha!h_\alpha^2$. The first directional derivative of $\mathcal{D}_{\max}$ is $2\max_j a_j/g_n$. At second order, take the largest individual second coefficient among the indices with maximal a_j. This accounts for the possible nondifferentiability of the maximum.

14.2 The stable tangent cone and the sharp local theorem

A *contrast vector* is a real vector orthogonal to $\mathbf{1}$. Let $P=I-\mathbf{1}\mathbf{1}^{\mathsf T}/n$ denote the projection onto the contrast subspace.

Theorem 14.3. *The possible first variations of normalized balanced stable polynomials at p_0 are exactly*

$$h(x)=-R(x)^{n-2}x^{\mathsf T}Bx,\qquad B=B^{\mathsf T}\succeq 0,\qquad B\mathbf{1}=0. \tag{156}$$

The matrix B is unique. Every such variation is realized by the straight path $p_t=p_0+th$ for all sufficiently small $t\geq 0$.

Proof. Fix a contrast vector x. For a one-sided differentiable stable path $p_t = p_0 + th + o(t)$, the polynomial $f_t(z) = p_t(z\mathbf{1} + x)$ is monic and real-rooted. Its coefficient of z^{n-1} is $\nabla p_t(\mathbf{1}) \cdot x = 0$. If its roots are $\rho_1(t), \dots, \rho_n(t)$, then

$$\sum_i \rho_i(t) = 0, \qquad [z^{n-2}]f_t = -\frac{1}{2}\sum_i \rho_i(t)^2 = O(t).$$

Thus each root is $O(\sqrt{t})$. Every elementary symmetric function of order $r \geq 3$ is $O(t^{r/2}) = o(t)$. The first variation therefore has only contrast degree two. Its quadratic form is nonpositive on the contrast subspace. Normalization removes contrast degree zero, and balance removes contrast degree one. This gives the unique matrix in (156). The same argument applies to limits of normalized coefficient differences along sequences.

Conversely, consider

$$p_t = R^{n-2}(R^2 - tx^{\mathsf{T}}Bx).$$

For small $t \geq 0$, the quadratic factor has strictly positive coefficients. Its matrix is $\mathbf{1}\mathbf{1}^{\mathsf{T}}/n^2 - tB$. It has one positive eigenvalue and all other eigenvalues are nonpositive. It is positive on every strictly positive real vector. If it vanished at $u + iv$ with $v > 0$, its real and imaginary parts would give $Q(u) = Q(v) > 0$ and $Q(u, v) = 0$. The orthogonal complement of a positive vector for a quadratic form with one positive direction is negative semidefinite. This contradicts $Q(u) > 0$. The quadratic factor is stable, and so is its product with R^{n-2}. The identity $B\mathbf{1} = 0$ gives normalization and balance. □

When $\operatorname{rank} B \geq 2$, the quadratic factor has rank at least three for $t > 0$. It cannot factor into two real linear forms. Thus the local theorem includes polynomials that are not products of linear forms.

Proposition 14.4. *For the tangent* (156)*, put* $\kappa_n = ((n-1)/n)^{n-3}$*. Then*

$$a_j = \kappa_n B_{jj}, \tag{157}$$

$$N(h) = c_n \frac{2n}{n-1} \operatorname{tr}(B^2), \tag{158}$$

$$a(h) = c_n \frac{n}{n-1} \operatorname{tr} B. \tag{159}$$

Consequently,

$$\mathcal{D}_{\mathrm{av}}(p_t) = 1 + \frac{2n}{(n-1)^2} \operatorname{tr} B\, t + o(t), \tag{160}$$

$$\mathcal{D}_{\max}(p_t) = 1 + 2\left(\frac{n}{n-1}\right)^2 \max_j B_{jj}\, t + o(t), \tag{161}$$

$$\mathcal{S}(p_t) = 1 + \frac{2n}{n-1} \operatorname{tr} B\, t + o(t). \tag{162}$$

Proof. Set every variable except $x_j = z$ equal to one. Then

$$h(1, \dots, z, \dots, 1) = -B_{jj}\left(\frac{n-1+z}{n}\right)^{n-2}(z-1)^2.$$

Its linear coefficient is (157). For (158), use an orthonormal coordinate with first coordinate $n^{-1/2}\sum_i x_i$. The remaining coordinates are contrast coordinates. The identity $\langle f, g\rangle_F = f(\partial)g|_{x=0}$ shows that the Fischer inner product is invariant under orthogonal changes of variables. The squared Fischer norm of a quadratic form with matrix B is $2\operatorname{tr}(B^2)$. The factor

R^{n-2} therefore gives $N(h) = 2(n-2)!n^{-(n-2)}\operatorname{tr}(B^2)$. This is (158). In the central coefficient, the diagonal terms of $x^{\mathsf{T}}Bx$ make no contribution. The off-diagonal entries sum to $-\operatorname{tr} B$ because $B\mathbf{1} = 0$. Thus $a(h) = (n-2)!n^{-(n-2)}\operatorname{tr} B$. The three expansions now follow from Lemma 14.2. □

Proposition 14.5. *Along the straight stable path in Theorem 14.3, the quadratic coefficient of $\mathcal{D}_{\mathrm{av}}$ is*

$$Q_n(B) = \frac{n^2(3n^2-5n+4)}{(n-1)^5}\sum_j B_{jj}^2 - \frac{2n(2n-1)}{(n-1)^3}\operatorname{tr}(B^2). \tag{163}$$

It is strictly negative for $n \geq 3$ and $B \neq 0$. The quadratic coefficient of $\mathcal{S}$ is

$$\frac{n^2}{(n-1)^2}(\operatorname{tr} B)^2 - \frac{2n}{n-1}\operatorname{tr}(B^2).$$

Proof. Assume first $n \geq 3$. Fix j, put $b = B_{jj}$, and let $u = B_{-j,j}$ and $C = B_{-j,-j}$. Write the remaining variables as $y = r\mathbf{1} + z$, with $\sum z_i = 0$. Let $r_j = u + (b/m)\mathbf{1}$ and $B_j = P_m C P_m$, where $P_m = I - \mathbf{1}\mathbf{1}^{\mathsf{T}}/m$. Coefficient extraction gives

$$s_j = \kappa_n\left(br^m - \frac{2}{n}r^{m-1}r_j\cdot z - \frac{n-2}{n}r^{m-2}z^{\mathsf{T}}B_j z\right), \qquad v_j = -\frac{2\kappa_n}{n}r_j.$$

The three contrast degrees are orthogonal for the Fischer inner product. Their squared norms, before multiplication by their displayed coefficients, are c_m, $c_m\|r_j\|^2$, and $2mc_m\operatorname{tr}(B_j^2)/(m-1)$. Lemma 14.2 gives the individual quadratic coefficient

$$\left(\frac{n}{m}\right)^4\left(b^2 - \frac{4}{n^2}\|r_j\|^2 + \frac{2m(n-2)}{n^2}\operatorname{tr}(B_j^2)\right) - \frac{2n}{m}\operatorname{tr}(B^2).$$

Expanding the projections and using $B\mathbf{1} = 0$ gives

$$\sum_j \|r_j\|^2 = \operatorname{tr}(B^2) - \frac{n}{n-1}\sum_j B_{jj}^2,$$

$$\sum_j \operatorname{tr}(B_j^2) = \frac{n(n-3)}{n-1}\operatorname{tr}(B^2) + \frac{n^2}{(n-1)^2}\sum_j B_{jj}^2.$$

Averaging proves (163). For $n = 2$, write $B = b\left(\begin{smallmatrix}1 & -1\\ -1 & 1\end{smallmatrix}\right)$ and extract the two linear deletions directly. The quadratic coefficient is zero, as (163) states. For every n, the zero row sums and Cauchy–Schwarz give $\sum_j B_{jj}^2 \leq (n-1)\operatorname{tr}(B^2)/n$. For $n \geq 3$, substitution yields

$$Q_n(B) \leq -\frac{n(n-2)(n+1)}{(n-1)^4}\operatorname{tr}(B^2) < 0.$$

Finally, insert (158) and (159) into the ratio defining $\mathcal{S}$. □

Theorem 14.6. *For each $n \geq 2$, there is a $\delta_n > 0$ such that every normalized balanced stable p with $0 < d_F(p, p_0) < \delta_n$ satisfies*

$$\mathcal{D}_{\mathrm{av}}(p) > 1, \qquad \mathcal{D}_{\max}(p) > 1, \qquad a(p)^2 > c_n N(p).$$

Moreover, the sharp asymptotic bounds are

$$\liminf_{p\to p_0}\frac{\mathcal{D}_{\mathrm{av}}(p)-1}{d_F(p,p_0)} \geq \frac{\sqrt{2n}}{(n-1)^{3/2}}, \tag{164}$$

$$\liminf_{p\to p_0}\frac{\mathcal{S}(p)-1}{d_F(p,p_0)} \geq \sqrt{\frac{2n}{n-1}}. \tag{165}$$

The limits run through distinct normalized balanced stable polynomials.

Proof. For a fixed nonzero tangent, divide (160) and (162) by the distance from (158). The resulting factors are the constants in (164) and (165), multiplied by $\operatorname{tr} B/\sqrt{\operatorname{tr}(B^2)} \geq 1$.

To include every approach to p_0, take an arbitrary sequence $p_\ell \to p_0$ and set $t_\ell = \sqrt{N(p_\ell - p_0)}$ and $h_\ell = (p_\ell - p_0)/t_\ell$. The coefficient unit sphere is compact. Each convergent subsequence of h_ℓ has a limit of norm one. The root-variance proof of Theorem 14.3, with $O(t_\ell)$ in place of $O(t)$, puts this limit in the same tangent cone. Analyticity gives a uniform first-order remainder on the compact sphere. Thus both lower bounds hold for the arbitrary sequence. Their positive constants give the stated neighborhood. Also $\mathcal{D}_{\max} \geq \mathcal{D}_{\mathrm{av}}$. Rank-one matrices $B = vv^{\mathsf{T}}$ with $v \perp \mathbf{1}$ realize equality in both asymptotic constants. □

The negative second coefficient in Proposition 14.5 is consistent with this theorem. The positive first coefficient dominates it near p_0. It also shows that convexity of the deletion ratio along these stable straight paths cannot prove a global inequality.

14.3 A variance-deficit quantity and exact local control

Proposition 14.7. *For every normalized balanced stable p, let*

$$B(p) = -\tfrac{1}{2} P \nabla^2 p(\mathbf{1}) P, \qquad E(p) = 2 \operatorname{tr} B(p).$$

Then $B(p) \succeq 0$ and

$$E(p) = n - 1 - \operatorname{tr} \operatorname{Cov}(\alpha) \geq 0, \tag{166}$$

where $\Pr(\alpha = \gamma) = p_\gamma$. Equality holds exactly at $p = p_0$. For a doubly stochastic matrix,

$$B(p_A) = \tfrac{1}{2}(A - J_n)^{\mathsf{T}}(A - J_n), \qquad E(p_A) = \|A - J_n\|_F^2.$$

Uniform variable mixing satisfies

$$p_s(x) = p((J_n + s(I - J_n))x), \qquad E(p_s) = s^2 E(p).$$

As $E(p) \to 0$ within this stable class,

$$\mathcal{D}_{\mathrm{av}}(p) = 1 + \frac{n}{(n-1)^2} E(p) + O_n(E(p)^{3/2}), \tag{167}$$

$$\mathcal{S}(p) = 1 + \frac{n}{n-1} E(p) + O_n(E(p)^{3/2}). \tag{168}$$

Proof. For a contrast vector x, the monic real-rooted polynomial $p(z\mathbf{1} + x)$ has root sum zero. Its quadratic contrast coefficient is therefore $-\frac{1}{2}\sum_i \rho_i^2 \leq 0$. This proves $B(p) \succeq 0$. Balance gives $\mathbb{E}\alpha = \mathbf{1}$ and

$$\operatorname{Cov}(\alpha) = \nabla^2 p(\mathbf{1}) + I - \mathbf{1}\mathbf{1}^{\mathsf{T}}.$$

This covariance annihilates $\mathbf{1}$ because $\sum_i \alpha_i = n$. Take the trace after projection to obtain (166). If $E = 0$, the positive semidefinite matrix $B(p)$ vanishes. Every contrast restriction has zero root variance and hence equals z^n. These restrictions determine p, so $p = p_0$. For p_A, direct differentiation gives $\nabla^2 p_A(\mathbf{1}) = \mathbf{1}\mathbf{1}^{\mathsf{T}} - A^{\mathsf{T}}A$. Projection proves the matrix formulas. The mixing map acts as multiplication by s on contrast vectors, which proves its exact scaling relation.

Expand by contrast degree,

$$p = R^n - R^{n-2} x^{\mathsf{T}} B(p) x + \sum_{r=3}^{n} R^{n-r} f_r(x),$$

where each f_r depends only on Px. For a unit contrast vector x, its restriction coefficients are elementary symmetric functions of the real roots. The elementary symmetric mean inequality, followed by Cauchy–Schwarz, gives

$$|f_r(x)| \le \binom{n}{r}\left(\frac{\sum_i \rho_i^2}{n}\right)^{r/2} = \binom{n}{r}\left(\frac{2x^T B(p)x}{n}\right)^{r/2}.$$

Finite-dimensional norm equivalence gives $f_r = O_n(E^{r/2})$ in coefficient norm. The quadratic contrast part is $O_n(E)$. Apply the analytic first-variation formulas to this expansion. The higher contrast degrees give $O_n(E^{3/2})$, and the analytic quadratic remainder is $O_n(E^2)$. This proves (167) and (168). □

Corollary 14.8. *For each $n \ge 2$, there is an $\varepsilon_n > 0$ such that every doubly stochastic A with $0 < \|A - J_n\|_F < \varepsilon_n$ satisfies both strict strong Gram inequalities and $\Delta_n(A) > 0$.*

Proof. The normalized balanced stable polynomials p_A and p_{A^T} depend continuously on A and converge to p_0 as $A \to J_n$. Proposition 14.7 shows that neither equals p_0 when $A \ne J_n$. Apply Theorem 14.6 to both polynomials and then Lemma 2.3. □

The quantity E does not make the deletion ratio monotone. Let $A = (I + P_3)/2$, where P_3 is a cyclic permutation, and let $A_s = J_3 + s(A - J_3)$. The central coefficient of p_{A_s} is $b(s) = 2/9 + s^2/12 - s^3/18$. Each of its six coefficients of type $(2,1,0)$ is $a(s) = 1/9 + s^3/72$. Every deletion is balanced, with $H(s) = b(s)^2 + 4a(s)^2$. Direct differentiation at $s = 1$ gives

$$H = \tfrac{1}{8}, \quad H' = \tfrac{1}{24}, \quad N = \tfrac{1}{4}, \quad N' = \tfrac{1}{8}, \qquad \frac{d}{ds}\mathcal{D}_{\mathrm{av}}(p_{A_s})\Big|_{s=1} = -\frac{3}{16}.$$

Thus the ratio decreases near this endpoint while E increases.

14.4 The Lorentzian local theorem

Theorem 14.9. *For each $n \ge 2$, a neighborhood of p_0 exists in which every distinct normalized balanced Lorentzian polynomial satisfies*

$$\mathcal{D}_{\mathrm{av}}(p) > 1 \qquad \textit{and} \qquad a(p)^2 > c_n N(p).$$

No sharp distance constant for this larger class is asserted.

Proof. At p_0, the Hessian of every order-$(n-2)$ derivative is $c_n \mathbf{1}\mathbf{1}^\mathsf{T}$. The normalized balanced Lorentzian tangent cone therefore consists exactly of the polynomials h with $h(\mathbf{1}) = 0$, $\nabla h(\mathbf{1}) = 0$, and

$$P\,\mathrm{Hess}(\partial^\beta h)P \preceq 0 \qquad (|\beta| = n-2). \tag{169}$$

Necessity follows by testing quadratic forms on the nullspace of the uniform Hessian. For sufficiency, decompose each Hessian along $\mathbf{1}$ and $\mathbf{1}^\perp$. Its Schur complement for $p_0 + th$ has the form $tC - t^2 bb^\mathsf{T}/(a_0 + ta)$, with $C \preceq 0$ and $a_0 > 0$. This is negative semidefinite for small $t \ge 0$. The coefficients remain positive, so full support and the Lorentzian characterization prove sufficiency.

Fix j and write

$$h(1,\dots,z,\dots,1) = \sum_{r=0}^{n} h_{j,r} z^r, \qquad b_r = \binom{n}{r}\frac{(n-1)^{n-r}}{n^n}, \qquad z_r = h_{j,r}/b_r.$$

The binary marginal, with the other variables equal, turns (169) into $2z_r - z_{r-1} - z_{r+1} \geq 0$ for $1 \leq r < n$. To see the sign directly, apply the mixed derivatives of that marginal to the binary contrast direction $(n-1, -1)$. Each resulting quadratic form is a positive sum of the forms in (169) evaluated at the contrast vector with entry $n-1$ in position j and -1 elsewhere. Its coefficients, divided by the uniform binomial coefficients, give the stated second difference. Thus (z_r) is concave. If $K \sim \mathrm{Bin}(n, 1/n)$, normalization and balance imply

$$\mathbb{E}z_K = 0, \qquad \mathbb{E}(Kz_K) = 0, \qquad \mathbb{E}K = 1.$$

Jensen's inequality gives $z_1 \geq 0$ and hence $a_j = g_n z_1 \geq 0$. If $a_j = 0$, equality in Jensen, together with the full support of K, forces (z_r) to be affine. The two moment constraints then force every z_r to vanish.

Suppose all a_j vanish. All one-coordinate marginals of h vanish, so the diagonal entries of $\nabla^2 h(\mathbf{1})$ are zero. Also $\nabla^2 h(\mathbf{1})\mathbf{1} = (n-1)\nabla h(\mathbf{1}) = 0$. The coefficient identity

$$\nabla^2 h(\mathbf{1}) = \sum_{|\beta|=n-2} \frac{\mathrm{Hess}(\partial^\beta h)}{\beta!}$$

makes its projection a positive weighted sum of the negative semidefinite matrices in (169). Its trace is zero, so it vanishes, and every summand vanishes. All second contrast derivatives of h are therefore zero polynomials. Hence $h = aR^n + R^{n-1}L$, where L is a contrast linear form. Normalization and balance give $a = L = 0$. The preceding argument proves $\sum_j a_j > 0$ for every nonzero tangent. Compactness on the coefficient unit sphere, followed by Lemma 14.2, gives the strict average-deletion neighborhood.

The norm statement follows by induction on n. For $n = 2$, a normalized balanced binary quadratic is $p = a(x^2 + y^2) + bxy$ with $2a + b = 1$. The Lorentzian condition gives $b \geq 2a$, and $b^2 - \frac{1}{2}N(p) = \frac{1}{2}(b^2 - 4a^2) \geq 0$, with equality only at p_0. For the induction step, each normalized balanced deletion $\widetilde{q}_j = q_j(D_j x)/q_j(D_j \mathbf{1})$ is Lorentzian and converges to the uniform power of degree m. Shrink the neighborhood until the induction hypothesis applies to every deletion. Since $a(q_j) = a(p)$, it gives $a(p)^2 \geq c_m H_j(p)$. Average these inequalities and use the strict deletion inequality to obtain $a(p)^2 \geq c_m n^{-1} \sum_j H_j(p) > c_m g_n N(p) = c_n N(p)$. □

14.5 Why balance alone does not suffice

Multiplication by x_j embeds $[x_j]p$ as the orthogonal projection onto the monomials with exponent exactly one in coordinate j. Thus

$$\sum_j N([x_j]p) = \sum_\alpha k_1(\alpha)\alpha! p_\alpha^2, \qquad \sum_j N(\partial_j p) = nN(p).$$

For the second identity, a monomial with exponent α contributes $\sum_j \alpha_j \alpha! p_\alpha^2 = n\alpha! p_\alpha^2$. The first operator has zero eigenvalues because some monomials have no exponent equal to one. It has no positive lower bound on the full coefficient space.

An exact perturbation also rules out the desired inequalities under positivity and balance alone. Let

$$V = (x_1 - x_2)(x_2 - x_3)(x_3 - x_1), \qquad p_\varepsilon = R^3 + \varepsilon V, \qquad 0 < |\varepsilon| < 1/9.$$

Every coefficient is positive, and normalization and balance hold. Direct coefficient calculation and binary-quadratic balancing give

$$a(p_\varepsilon) = \frac{2}{9}, \quad N(p_\varepsilon) = \frac{2}{9} + 12\varepsilon^2, \quad H_j(p_\varepsilon) = \frac{8}{81} - 4\varepsilon^2.$$

Therefore

$$\frac{1}{3}\sum_j H_j(p_\varepsilon) - g_3 N(p_\varepsilon) = -\frac{28}{3}\varepsilon^2 < 0,$$

and (150) fails as well. For a contrast x with $V(x) \neq 0$, the restriction is $z^3 + \varepsilon V(x)$. It has nonreal zeros, so this perturbation is not stable. Theorem 14.9 also excludes it from the Lorentzian class for sufficiently small nonzero ε.

14.6 Subtraction to a pure-coefficient boundary

Theorem 14.10. *Let p be a balanced Lorentzian polynomial of degree n in n variables with nonnegative coefficients. Put $S = \sum_i x_i$ and $\mu = \min_i [x_i^n]p$. Then $q = p - \mu S^n$ is nonnegative, balanced, and Lorentzian. It has a zero pure coefficient and may be zero. Consequently,* (150) *holds for the full balanced Lorentzian class if and only if it holds on the subclass with a zero pure coefficient.*

Proof. Let $d_\alpha = p_\alpha / \binom{n}{\alpha}$. The two-by-two principal minors of the derivative Hessians give

$$d_\alpha^2 \geq d_{\alpha - e_i + e_j} d_{\alpha + e_i - e_j} \qquad (\alpha_i, \alpha_j > 0).$$

First assume all coefficients are positive. Among the indices minimizing d_α, choose one maximizing $\sum_i \alpha_i^2$. If two coordinates are positive, the displayed inequality and minimality force both adjacent indices to attain the same minimum. At least one has a larger squared coordinate sum, a contradiction. Thus the minimum occurs at a pure index, and

$$p_\alpha \geq \mu \binom{n}{\alpha}. \tag{170}$$

For a nonzero polynomial with zero coefficients, substitute $((1-\varepsilon)I + \varepsilon J_n)x$ for x. This preserves balance and the Lorentzian property and makes every coefficient positive. Let $\varepsilon \downarrow 0$ in (170). The zero polynomial is immediate.

For $0 \leq t < \mu$, every coefficient of $p - tS^n$ is positive and

$$\operatorname{Hess}(\partial^\beta (p - tS^n)) = \operatorname{Hess}(\partial^\beta p) - tn!\mathbf{1}\mathbf{1}^{\mathsf{T}}.$$

Subtracting a positive semidefinite matrix cannot increase the number of positive eigenvalues. Full support therefore gives the Lorentzian property. Closedness gives the endpoint $t = \mu$. Balance is linear, and the coefficient attaining μ becomes zero.

The set $\{f : a(f) \geq \sqrt{c_n}\sqrt{N(f)}\}$ is a convex cone. Also $a(S^n) = n!$ and $N(S^n) = n!n^n$, so S^n belongs to its boundary. If q satisfies the target inequality, the triangle inequality gives

$$\sqrt{c_n N(p)} \leq \sqrt{c_n N(q)} + \mu\sqrt{c_n N(S^n)} \leq a(q) + \mu n! = a(p).$$

The reverse implication is inclusion of the boundary subclass. □

There is also an exact ratio statement. For normalized nonnegative q, put $p_\lambda = (1-\lambda)p_0 + \lambda q$, $u = a(q)/c_n - 1$, and $v = N(q)/c_n - 1 \geq 0$. Orthogonality gives

$$\mathcal{S}(p_\lambda) = \frac{(1+\lambda u)^2}{1 + \lambda^2 v} \geq \min\{1, \mathcal{S}(q)\}, \qquad 0 \leq \lambda \leq 1. \tag{171}$$

Indeed, its derivative has the sign of $u - \lambda v$ whenever $1 + \lambda u > 0$. An interior critical point can only be a maximum. The endpoints with zero numerator follow by continuity. If a normalized nonuniform p has $\mu > 0$, then

$$q = \frac{p - \mu S^n}{1 - \mu n^n}, \qquad p = (1 - \lambda)p_0 + \lambda q, \qquad \lambda = 1 - \mu n^n.$$

The denominator is positive by (170) and nonuniformity. Thus a counterexample can be moved to the pure-coefficient boundary with no increase in its ratio. Along any admissible subtraction direction, a ratio below one strictly decreases as λ increases. There is no interior local minimum below one.

This reduction does not connect the boundary to the local theorem through admissible line segments. For example, $q = \frac{1}{4}x_1(x_2 + x_3)^2$ is normalized, balanced, and stable. For every $0 < \lambda < 1$, the Hessian of $\partial_1((1 - \lambda)R^3 + \lambda q)$ is a positive sum of two nonparallel positive rank-one matrices. It has two positive eigenvalues. Thus the intervening polynomial is not Lorentzian.

14.7 Cubic stability, deletion, and boundary rigidity

Proposition 14.11. *For a balanced nonnegative stable ternary cubic, the subtraction in Theorem 14.10 preserves stability, with zero permitted at the endpoint.*

Proof. For a contrast vector y, write $p(s\mathbf{1}+y) = as^3+u(y)s+v(y)$, where $a = p(\mathbf{1}) > 0$. Balance removes the quadratic coefficient. Real-rootedness gives $u(y) \le 0$ and $-4u(y)^3 \ge 27av(y)^2$. Subtracting tS^3 changes a to $a_t = a - 27t$ and leaves u, v fixed. If $a_t > 0$, the same discriminant inequality holds with a_t. Thus the residual is hyperbolic in direction $\mathbf{1}$. Its coefficients remain nonnegative by Theorem 14.10, so it is stable. Closure handles the endpoint. If $a_t = 0$, nonnegative coefficients and zero value at $\mathbf{1}$ force the residual to be zero. □

For a ternary cubic, write its pure coefficients as U_i, its central coefficient as b, and its deletions as $[x_j]p = a_j x_k^2 + b x_k x_\ell + c_j x_\ell^2$. Binary balancing gives $H_j = b^2 + 4a_j c_j$. If one of a_j, c_j is zero, this formula denotes the limit along positive diagonals and equals the corresponding infimum. A coefficient calculation gives

$$G(p) := 3\sum_j H_j - 4N(p) = 5b^2 - \sum_j \left((a_j + c_j)^2 + 7(a_j - c_j)^2\right) - 24\sum_i U_i^2. \tag{172}$$

The inequality $G \ge 0$ is precisely the average-deletion inequality in degree three. For $b \ge 0$, it defines a convex cone because its right side is five times a linear coordinate squared minus a sum of squares. The polynomial S^3 lies on its boundary. Theorem 14.10 therefore reduces the Lorentzian cubic deletion problem to zero pure coefficients. Proposition 14.11 makes the same reduction inside the stable cubic class. The deletion inequality on that full boundary subclass remains unproved here.

A proposed estimate using only the mixed coefficients would require $\sum_j((a_j+c_j)^2+7(a_j-c_j)^2) \le 3b^2$. It fails in the Lorentzian class. Consider

$$p(x, y, z) = (y + z)\left(\frac{(y + z)^2}{4\sqrt{3}} + \frac{x(y + z)}{2} + \frac{x^2}{\sqrt{3}}\right).$$

It is balanced, with $p(\mathbf{1}) = 2 + 4\sqrt{3}/3$, and $b = 1$, $U_x = 0$. Its y and z derivatives are squares of linear forms. Its x derivative is $(y + z)((y + z)/2 + 2x/\sqrt{3})$. These quadratics have at most

one positive Hessian eigenvalue. Its support is the degree-three simplex without $(3,0,0)$, which satisfies the exchange property. Thus it is Lorentzian. Direct expansion gives

$$N(p) = 13/3, \quad H_x = H_y = H_z = 2, \quad G(p) = 2/3, \quad \sum_j ((a_j + c_j)^2 + 7(a_j - c_j)^2) = 10/3.$$

The quadratic factor, viewed in x and $y + z$, has discriminant $-1/12$ and is not stable. This example does not disprove the proposed estimate restricted to stable cubics. That restricted estimate is also unproved here.

Proposition 14.12. *Let $p(x, y, z)$ be a nonzero balanced nonnegative stable cubic with zero x^3 coefficient. If $\mathrm{Hess}(\partial_x p)$ is singular, then p factors over the reals.*

Proof. If the degree in x is at most one, balance says that the mean x-exponent is one. Every nonzero monomial must therefore have x-exponent one, and $p = xq(y, z)$. Otherwise the degree in x is two. The derivative $\partial_x p$ is a stable quadratic with singular Hessian and zero x^2 coefficient. Its rank cannot be one because a rank-one quadratic with that coefficient zero is independent of x. Its rank is therefore two, with one positive and one negative eigenvalue, so it factors into real linear forms. The stable linear factors can be chosen with nonnegative coefficients. Exactly one depends on x. Thus $\partial_x p = L(y, z)(cx + M(y, z))$, with $c > 0$ and nonnegative linear forms L, M. Integration gives

$$p = L\left(\frac{c}{2}x^2 + Mx\right) + r(y, z).$$

For every real y, z with $L \neq 0$, specialization gives a nonzero stable quadratic in x. Its discriminant is $L(LM^2 - 2cr) \geq 0$. If r did not vanish on the line $L = 0$, crossing that line near a point where $r \neq 0$ would change the sign of this discriminant. Therefore L divides r, and $p = LQ$. □

Subtraction of a general balanced stable polynomial does not preserve the Lorentzian property. For an exact example, take

$$p = (2y + 4z)(2x + 3y + z)(4x + y + z), \qquad h = \tfrac{1}{8}(y + z)(2x + y + z)^2.$$

Both are balanced stable products with zero x^3 coefficient. Their values at $\mathbf{1}$ are 216 and 4. In the order $(x^2y, x^2z, xy^2, xyz, xz^2, y^3, y^2z, yz^2, z^3)$ their coefficients are

$$(16, 32, 28, 68, 24, 6, 20, 18, 4), \qquad (\tfrac{1}{2}, \tfrac{1}{2}, \tfrac{1}{2}, 1, \tfrac{1}{2}, \tfrac{1}{8}, \tfrac{3}{8}, \tfrac{3}{8}, \tfrac{1}{8}).$$

All displayed coefficients of $p - h$ are positive, but

$$\mathrm{Hess}(\partial_x(p - h)) = \begin{pmatrix} 0 & 31 & 63 \\ 31 & 55 & 67 \\ 63 & 67 & 47 \end{pmatrix}$$

has determinant -1760 and trace 102. It has two positive eigenvalues, so $p - h$ is not Lorentzian. The failure occurs for arbitrarily small positive subtraction as well. For $v = (-11, 16, -8)^{\mathsf{T}}$, $\mathrm{Hess}(\partial_x p)v = 0$ and $v^{\mathsf{T}} \mathrm{Hess}(\partial_x h)v = -112$. The zero eigenvalue therefore becomes positive under small subtraction, while the existing positive eigenvalue remains positive. This prevents an unrestricted iteration of the uniform-power subtraction argument.

14.8 An admissible attenuation and a finite support identity

Proposition 14.13. *Let p be Lorentzian, and let r be its maximum degree in variable x_i. Multiply each coefficient with exponent $\alpha_i = r$ by $\lambda \in [0,1]$. The resulting polynomial p_λ is Lorentzian.*

Proof. For $\lambda > 0$, the support is unchanged and is M-convex. Consider a quadratic derivative $\partial^\beta p$. If $\beta_i = r-2$, attenuation decreases only its ii Hessian entry. This subtracts a positive semidefinite diagonal matrix. If $\beta_i = r-1$, the ii entry is zero and the changed row and column are obtained by diagonal congruence with factor λ in position i. If $\beta_i = r$, the Hessian is multiplied by λ and its ith row and column are zero. All other Hessians are unchanged or zero. Each retains at most one positive eigenvalue. The characterization proves the assertion for $\lambda > 0$, and closedness gives $\lambda = 0$. □

This operation permits positive diagonal balancing whenever the mean exponent is in the relative interior of the support polytope. It does not imply monotonicity or convexity of the balanced squared Fischer norm.

For completeness, one support class has an exact entropy reformulation. Let S be a nonempty M-convex set of degree-n exponent vectors. Assume $\mathbf{1} \in S$ and that $\mathbf{1}$ belongs to the relative interior of $\operatorname{conv} S$. Define

$$p_S(x) = \sum_{\alpha \in S} \frac{x^\alpha}{\alpha!}, \qquad Q_\alpha = \frac{c_n}{\alpha!}, \qquad F_S(z) = \sum_{\alpha \in S} Q_\alpha e^{(\alpha - \mathbf{1})\cdot z}.$$

Here Q is the multinomial probability distribution on all degree-n exponent vectors. The polynomial p_S is Lorentzian by Brändén and Huh [3, Theorem 3.10]. On the span of the vectors $\alpha - \mathbf{1}$, the relative-interior assumption makes F_S coercive and strictly convex. Indeed, every nonzero direction has a positive inner product with some $\alpha - \mathbf{1}$, and its Hessian is a positive weighted sum of their outer products. It has a unique minimizer modulo orthogonal redundancies. Choose a representative z_* with $\sum_i z_{*,i} = 0$. Then $\nabla F_S(z_*) = 0$.

The tilted polynomial $q(x) = p_S(e^{z_{*,1}}x_1, \dots, e^{z_{*,n}}x_n)$ is balanced and has central coefficient one. Direct substitution gives

$$c_n N(q) = F_S(2z_*). \tag{173}$$

Thus its target inequality is exactly $F_S(2z_*) \le 1$. This last inequality is not proved for arbitrary S here.

To state the entropy identity, put

$$P_*(\alpha) = \frac{Q_\alpha \mathbf{1}_{\{\alpha \in S\}} e^{(\alpha - \mathbf{1})\cdot z_*}}{F_S(z_*)}.$$

This distribution has mean $\mathbf{1}$. For probability distributions P, Q with $Q > 0$, define $D(P\|Q) = \sum_\alpha P_\alpha \log(P_\alpha/Q_\alpha)$ and $D_2(P\|Q) = \log \sum_\alpha P_\alpha^2/Q_\alpha$. These are the relative entropy and the order-two Rényi divergence. The zero mean of $\alpha - \mathbf{1}$ gives

$$D(P_*\|Q) = -\log F_S(z_*) = \log \frac{P_*(\mathbf{1})}{Q_{\mathbf{1}}}, \qquad D_2(P_*\|Q) = \log F_S(2z_*) - 2\log F_S(z_*).$$

Therefore (173) makes the remaining inequality equivalent to $D_2(P_*\|Q) \le 2D(P_*\|Q)$. These are exact identities, not a proof of that entropy inequality. If the mean lies on a proper face, the same identities apply after restriction to the minimal face containing it, provided the central exponent remains in that face. The interior argument then takes place in its affine span.

14.9 A multivariate obstruction to two-sided scalar localization

The scalar localization proof uses two-sided linear coefficient perturbations. At an active binary log-concavity equality, its linear constraint makes the quadratic error nonnegative. This implication fails for multivariate derivative Hessians. The following example also shows that binary coefficient inequalities and coordinate marginals do not replace Lorentzianity.

First consider a real symmetric matrix with one positive eigenvalue, r negative eigenvalues, and a nontrivial kernel. In orthonormal range and kernel coordinates, write it and a perturbation as

$$H = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \qquad K = \begin{pmatrix} A & B \\ B^{\mathsf{T}} & C \end{pmatrix},$$

where D is invertible and has inertia $(1, r)$. For sufficiently small ε, the range block $D + \varepsilon A$ retains that inertia. Its kernel Schur complement in $H + \varepsilon K$ is

$$\varepsilon C - \varepsilon^2 B^{\mathsf{T}}(D + \varepsilon A)^{-1} B.$$

If C has a nonzero eigenvalue, one sign of ε makes this complement positive in a direction. That direction supplies a second positive eigenvalue of the full matrix. Thus preservation of at most one positive eigenvalue for both signs requires $C = 0$. Under this necessary condition, the exact remaining condition is

$$B^{\mathsf{T}}(D + \varepsilon A)^{-1} B \succeq 0. \tag{174}$$

In particular, $B^{\mathsf{T}} D^{-1} B \succeq 0$ is necessary. It is sufficient for both small signs when $\ker(B^{\mathsf{T}} D^{-1} B) = \ker B$. Indeed, the form is positive definite on the orthogonal complement of $\ker B$, remains positive there by continuity, and vanishes on $\ker B$ for every ε. The simpler condition $K(\ker H) = 0$ is also sufficient. When D has only a positive eigenvalue, these conditions explain the favorable binary behavior. When D also has a negative direction, the linear condition $C = 0$ does not control the sign in (174).

Now set

$$p_0 = (x + y)(y + z)(z + x) = x^2 y + x^2 z + xy^2 + xz^2 + y^2 z + yz^2 + 2xyz.$$

This is a stable, and hence Lorentzian, product polynomial. It is balanced because $p_0(\mathbf{1}) = 8$ and each first derivative at $\mathbf{1}$ is 8. Moreover,

$$\frac{p_0}{8} = p_A, \qquad A = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \Omega_3.$$

The polynomial cannot be expressed in one variable and the sum of the other two variables. For example, at $x = 0$ it equals $yz(y + z)$, which is not determined by $y + z$ alone. Symmetry gives the same conclusion for the other choices of special variable.

Suppose $p_0 \pm \varepsilon h$ is nonnegative, balanced, Lorentzian, and has central coefficient 2 for both signs of every sufficiently small $\varepsilon > 0$. The zero pure coefficients of p_0 force the pure coefficients of h to vanish. Its central coefficient also vanishes. Thus

$$h = A_1 x^2 y + B_1 x^2 z + C_1 xy^2 + D_1 xz^2 + E_1 y^2 z + F_1 yz^2.$$

The kernels of the three derivative Hessians of p_0 are spanned, respectively, by $(0, 1, -1)$, $(1, 0, -1)$, and $(1, -1, 0)$. The necessary zero kernel-compression conditions give

$$C_1 + D_1 = 0, \qquad A_1 + F_1 = 0, \qquad B_1 + E_1 = 0.$$

They imply $h(\mathbf{1}) = 0$. The balance equations for x and y then give $A_1+B_1 = 0$ and $C_1-B_1 = 0$. Consequently every possible direction is a scalar multiple of

$$V = (x-y)(x-z)(y-z), \qquad (A_1, B_1, C_1, D_1, E_1, F_1) = a(1,-1,-1,1,1,-1).$$

For $p_t = p_0 + tV$, direct differentiation gives

$$H_x(t) = \mathrm{Hess}(\partial_x p_t) = \begin{pmatrix} 0 & 2(1+t) & 2(1-t) \\ 2(1+t) & 2(1-t) & 2 \\ 2(1-t) & 2 & 2(1+t) \end{pmatrix},$$

and

$$\det H_x(t) = -64t^2, \qquad \operatorname{tr} H_x(t) = 4. \tag{175}$$

For $t \neq 0$, the negative determinant and positive trace force exactly two positive eigenvalues and one negative eigenvalue. Thus no nonzero supported two-sided direction is feasible. The other two derivative Hessians have the same determinant and trace. In the basis e_1, e_2+e_3, e_2-e_3, the displayed Hessian is congruent to

$$\begin{pmatrix} 0 & 4 & 4t \\ 4 & 8 & -4t \\ 4t & -4t & 0 \end{pmatrix}.$$

Its range block is $D = \left(\begin{smallmatrix} 0 & 4 \\ 4 & 8 \end{smallmatrix}\right)$ and its coupling vector is $b = (4,-4)^{\mathsf{T}}$. Since $b^{\mathsf{T}} D^{-1} b = -16$, the kernel Schur complement is $16t^2$. This gives the exact quadratic failure that the scalar dimension count misses.

For $|t| < 1$, all seven supported coefficients of p_t are positive. The polynomial remains balanced and has central coefficient 2. Its three coordinate marginals are unchanged:

$$p_t(x,1,1) = 2 + 4x + 2x^2,$$

with the analogous identities in y and z. Each 2×2 principal minor of each derivative Hessian is one of

$$-4(1+t)^2, \qquad -4(1-t)^2, \qquad -4t^2.$$

Thus every binary coefficient-line log-concavity condition holds. The coordinate probability marginals are also ultra-log-concave. Nevertheless, the full Hessian condition fails by (175). The squared Fischer norm is

$$N(p_t) = 16 + 12t^2.$$

At $t = 1/2$, the normalized polynomial $P = p_t/8$ satisfies

$$[xyz]P = \frac{1}{4}, \qquad N(P) = \frac{19}{64}, \qquad c_3 N(P) - ([xyz]P)^2 = \frac{1}{288} > 0.$$

This refutes an extension of the norm bound based only on balance, marginal ultra-log-concavity, and binary coefficient inequalities. It is neither a Lorentzian counterexample nor a Gyires counterexample.

The rigidity assertion concerns two-sided linear perturbations on the original support. It does not make p_0 a global norm maximizer at fixed central coefficient. Indeed, consider the one-sided path

$$q_s = p_0 + \frac{s}{3}(x^3 + y^3 + z^3), \qquad 0 \le s \le 1.$$

It remains balanced and has central coefficient 2. The derivative Hessians have a zero contrast direction. On the complementary two-dimensional subspace, their congruent block is $\left(\begin{smallmatrix} 2s & 4 \\ 4 & 8 \end{smallmatrix}\right)$,

with determinant $16(s-1) \le 0$ and positive trace. Their full signatures therefore have at most one positive eigenvalue. For $s > 0$ the coefficient support is the full degree-three simplex; at $s = 0$ the polynomial is the stable product already considered. The Lorentzian criterion thus applies along the whole path. Also,

$$N(q_s) = 16 + 2s^2, \qquad q_1 = \frac{(x+y+z)^3}{3}, \qquad \frac{([xyz]p_0)^2}{c_3 N(p_0)} = \frac{9}{8}.$$

The example excludes the direct claim that every non-binary input can be removed by a supported two-sided perturbation. A broader localization proof must also treat changes of support or compare the norms at other rigid points.

15 Exact expectations in a regular bipartite configuration model

Let $d \ge 2$ and $n \ge 1$. Give each of n row vertices and n column vertices exactly d labeled stubs, and choose a uniformly random bijection between the dn row stubs and the dn column stubs. Let G be the resulting $n \times n$ nonnegative integer adjacency matrix, with parallel edges counted. Then G/d is doubly stochastic in every realization.

Define

$$F_d(x,y) = \sum_{\substack{k,r \ge 0 \\ k+2r \le d}} \frac{d!(k+r)!}{k!(r!)^2(d-k-2r)!}\, x^k y^r.$$

The expected Gram permanent is

$$\boxed{E\,\mathrm{per}(GG^T) = \frac{n!d^n}{(dn)!} \sum_{t=0}^{n} (d-1)^{n-t}((d-2)n+t)!(n-t)![x^t y^{n-t}]F_d(x,y)^n.} \tag{176}$$

15.1 Counting proof

Expand the Gram permanent as a pair of edge choices, colored red and blue, with one red and one blue edge at every row. At a column, the red edges and blue edges are paired; if there are ℓ of each color, this contributes $\ell!$.

Let t be the number of row vertices that use the same edge for both colors. These contribute one stub of type RB each; every other row contributes distinct R-only and B-only stubs. The number of ways to color the row stubs is

$$\binom{n}{t} d^t \big(d(d-1)\big)^{n-t}.$$

At a column, suppose k stubs are RB, r are R-only, and r are B-only. The remaining $d-k-2r$ stubs are unused. There are

$$\frac{d!}{k!(r!)^2(d-k-2r)!}$$

such stub colorings and $(k+r)!$ pairings of the colored edges, producing the corresponding term of F_d. Across all columns, the prescribed color totals are $t, n-t, n-t, (d-2)n+t$.

For any fixed row and column colorings with these totals, the probability that a uniform stub bijection respects all four colors is

$$\frac{t!\,(n-t)!^2\,((d-2)n+t)!}{(dn)!}.$$

Multiplying and cancelling factorials gives (176).

For illustration,

$$F_2(x,y) = 1 + 2x + 2x^2 + 2y,$$

$$F_3(x,y) = 1 + 3x + 6x^2 + 6x^3 + 6y + 12xy.$$

15.2 The second moment

For two perfect matchings, each column must receive exactly one red and one blue edge. Thus its local possibilities are RB or R-only plus B-only. Replace F_d in (176) by

$$L_d(x,y) = dx + d(d-1)y.$$

The result is exactly $E\,\mathrm{per}(G)^2$. In particular, if

$$w_t = (d-1)^{n-t}((d-2)n+t)!(n-t)!,$$

then

$$\frac{c_n E\,\mathrm{per}(GG^T)}{E\,\mathrm{per}(G)^2} = \frac{n!}{n^n}\frac{\sum_t w_t [x^t y^{n-t}] F_d(x,y)^n}{\sum_t w_t \binom{n}{t} d^n (d-1)^{n-t}}. \tag{177}$$

Every quantity in (177) is an explicitly computable integer before the final division.

These are exact finite identities. A violation of the ratio bound for one pair of parameters would imply a realization that violates the strong inequality. No such violation is established here. The identities do not assert the unrestricted Gyires conjecture.

16 Exact limits of stronger proof steps

The preceding proofs use specific matrix or polynomial structure. This section gives exact counterexamples to several stronger statements. Each counterexample concerns an auxiliary statement. None is a counterexample to Gyires's inequality. The squared Fischer norm remains $N(p) = \sum_\alpha \alpha! p_\alpha^2$. Put $g_n = ((n-1)/n)^{n-1} = c_n/c_{n-1}$. For a singleton deletion $q_j = [x_j]p$, let $H(q_j)$ be its squared Fischer norm after capacity balancing with determinant one. A balanced limit is used when a finite positive scaling does not exist. Thus, at a positive capacity-balancing diagonal matrix D,

$$H(q_j) = \frac{N(q_j(Dx))}{(\det D)^2}.$$

16.1 Negative coefficients in a general Birkhoff sorting chamber

The sorted-gap proof for circulants does not extend to arbitrary lists of permutation matrices by the same coefficient-sign assertion. Let E be the order-four all-ones matrix, and define

$$A(t,u,v) = 6tE + u(3E - 2I) + vI, \qquad t,u,v \geq 0.$$

The sum of all 24 order-four permutation matrices is $6E$. The identity and the nine derangements sum to $3E - 2I$. Indeed, each off-diagonal position occurs in three derangements. These subsets are nested, and the last subset can be the identity alone. Thus t, u, v are three gap variables in an actual sorting chamber for Birkhoff coefficients, with all other gaps set to zero. Nevertheless,

$$[t^6uv](4^4 \operatorname{per}(A)^2 - 4! \operatorname{per}(AA^{\mathsf{T}})) = -6879707136. \tag{178}$$

The coefficient can be computed without expanding the full permanent. Put $H = I - E/4$, so $\|H\|_F^2 = 3$. For an order-four matrix X with zero row and column sums,

$$\operatorname{per}(aE + X) = 24a^4 + a^2\|X\|_F^2 + O(X^3).$$

The linear term vanishes. For the quadratic term, summing products from two distinct rows and columns and using the zero sums gives $\|X\|_F^2/2$ before the factor $2!a^2$. Also

$$(aE + X)(aE + X)^{\mathsf{T}} = 4a^2E + XX^{\mathsf{T}}.$$

The term linear in XX^{T} in its permanent vanishes, since the sum of the entries of XX^{T} is zero. The quadratic part in X of the defect is therefore $256 \cdot 24 \cdot 2\, a^6\|X\|_F^2$. In the displayed family, $X = (-2u + v)H$. For the term of total degree two in u, v, the scalar part can be replaced by $a = 6t$. Its contribution is

$$256 \cdot 24 \cdot 2 \cdot 6^6 \cdot (-12) = -6879707136,$$

which proves (178). The exact full expansion has twelve negative coefficients. A negative coefficient does not imply a negative value on the nonnegative orthant. It excludes this coefficientwise proof step only.

16.2 A zero-containing column need not be a useful deletion

The condition $[x_j^n]p = 0$ does not imply $H([x_j]p) \geq g_n N(p)$. It also does not suffice to choose the best column among those with this vanishing coefficient. Both failures occur for products of doubly stochastic row forms.

First let $A = M/8$, where

$$M = \begin{pmatrix} 0 & 0 & 5 & 3 \\ 3 & 0 & 0 & 5 \\ 2 & 3 & 3 & 0 \\ 3 & 5 & 0 & 0 \end{pmatrix}, \qquad p_M(x) = \prod_{i=1}^{4} (Mx)_i.$$

Every row and column sum of M is eight. Working with M instead of A multiplies both squared norms by 8^8. It does not change their ratio. Direct coefficient expansion gives

$$N(p_M) = 2276500, \qquad q_1(y,z,w) = [x_1]p_M = (5z + 3w)(45y^2 + 45yz + 95yw + 45zw).$$

Let t be the unique positive root of

$$f(t) = 19t^4 + 9t^3 - 15t - 15 = 0.$$

Descartes' rule gives at most one positive root, and exact endpoint substitution gives

$$t \in I := \left[\frac{1021945}{10^6}, \frac{1021946}{10^6}\right].$$

The signs at the two endpoints are opposite, so this root exists. The scale $D = (t, 1, t^2)$ capacity-balances q_1. To check this, write its quadratic factor as Q. The equation $yQ_y = Q$ at D reduces to $45t^2 = 45t^2$. The expression $w(q_1)_w - q_1$ at D equals $15tf(t)$. Homogeneity gives the remaining balance equation. The function

$$(z_1, z_2, z_3) \longmapsto \log q_1(e^{z_1}, e^{z_2}, e^{z_3}) - z_1 - z_2 - z_3$$

is convex. Its stationary point is therefore a global capacity minimum. The support spans the degree-three affine hyperplane, so the scale is unique modulo common multiplication.

Expanding the norm at this scale gives

$$H(q_1) = 101250t^{-4} + 202500t^{-2} + 372100 + 72900t^2 + 162450t^4. \tag{179}$$

All coefficients are positive. Use the lower endpoint of I for negative powers and the upper endpoint for positive powers to obtain an upper bound. The resulting rational comparison gives

$$\frac{H(q_1)}{(27/64)N(p_M)} < \frac{95}{100} < 1.$$

This proves the prescribed-column failure. The corresponding lower interval bound also gives $H(q_1) > 911200 = N(q_1)$. Thus capacity balancing can increase this individual gauge-normalized norm.

For a failure with only one eligible column, use

$$A' = \frac{1}{7200}\begin{pmatrix} 0 & 24 & 4479 & 2697 \\ 2697 & 16 & 16 & 4471 \\ 1806 & 2689 & 2689 & 16 \\ 2697 & 4471 & 16 & 16 \end{pmatrix} = \frac{M'}{7200}. \tag{180}$$

This matrix is doubly stochastic, and its only zero entry is the first entry of the first row. Its product polynomial has only one zero degree-four coefficient, namely $[x_1^4]p_{M'}$. For any other monomial, assign one occurrence of another variable to the first row. All remaining rows have positive entries.

The following rational interval calculation establishes a capacity-balancing scale and bounds the resulting squared Fischer norm. Put $q = [x_1]p_{M'}$, and use the gauge $D = (y, z, 1)$. For $q = \sum_{a+b+c=3} q_{abc} y^a z^b w^c$, its ten coefficients, in the order indicated, are

(a,b,c)	q_{abc}	(a,b,c)	q_{abc}
$(0,0,3)$	870637556880	$(0,1,2)$	89561760156999
$(0,2,1)$	146964165788097	$(0,3,0)$	1041518854368
$(1,0,2)$	185659808813829	$(1,1,1)$	397840876407234
$(1,2,0)$	147375866123649	$(2,0,1)$	89762967708687
$(2,1,0)$	147118426392969	$(3,0,0)$	784079123688

and

$$N(p_{M'}) = 966483494962008824960673644628.$$

Define

$$F_1(y,z) = yq_y(y,z,1) - q(y,z,1), \qquad F_2(y,z) = zq_z(y,z,1) - q(y,z,1).$$

Take the rational box

$$X = \left[\frac{9789}{10000}, \frac{9791}{10000}\right] \times \left[\frac{9584}{10000}, \frac{9586}{10000}\right], \qquad c = \left(\frac{979}{1000}, \frac{1917}{2000}\right), \qquad R = F'(c)^{-1}.$$

The determinant defining R is nonzero. Evaluate the derivatives as interval polynomials on X to obtain an interval Jacobian $J(X)$. Ordinary rational interval arithmetic gives

$$c - RF(c) + (I - RJ(X))(X - c) \subset \left[\frac{9789782}{10^7}, \frac{9789784}{10^7}\right] \times \left[\frac{9584771}{10^7}, \frac{9584773}{10^7}\right] \subset \operatorname{int} X.$$

The mean-value formula therefore shows that $x \mapsto x - RF(x)$ maps X into itself. Brouwer's fixed-point theorem gives a zero of F in X. Convexity of the logarithmic capacity objective makes it a capacity minimum. Positivity of all ten coefficients gives uniqueness up to a common scale.

For every $(y, z) \in X$, evaluate

$$\frac{N(q(yY, zZ, W))}{(yz)^2} = \sum_{a+b+c=3} a!b!c!q_{abc}^2 y^{2a-2} z^{2b-2}$$

with rational interval arithmetic. The upper endpoint gives

$$\frac{H(q)}{(27/64)N(p_{M'})} < \frac{956}{1000} < 1. \tag{181}$$

The stated coefficients, box, and interval rules specify the certificate completely. The accompanying checker uses integers and rational numbers for every decisive comparison. Since column one is the only zero-containing column, (181) also excludes a maximum restricted to such columns. For comparison, direct integer permanent evaluation gives the row strong deficit $\operatorname{per}(M')^2 - c_4 \operatorname{per}(M'M'^{\mathsf{T}})$, the column strong deficit $\operatorname{per}(M')^2 - c_4 \operatorname{per}(M'^{\mathsf{T}}M')$, and the Gyires deficit $\Delta_4(M')$:

$$\begin{aligned} \text{Row strong deficit} &= \frac{5413562823023016913579550125 77}{8}, \\ \text{Column strong deficit} &= \frac{5409614616465032660547506335 05}{8}, \\ \text{Gyires deficit} &= 29909841945233088877175861 5932. \end{aligned}$$

All three are positive. No maximum over all columns is refuted.

16.3 The ultra-log-concave class does not satisfy the stronger contraction

Use the binomial-noise operator from the rank-two reduction. Set $n = 4$, $t = 2$, $B = \operatorname{Bin}(4, 1/2)$, and

$$P = \delta_2, \qquad Q = (0, 2/7, 3/7, 2/7, 0), \qquad r = P/B, \quad s = Q/B.$$

Both laws have mean two and log-concave likelihood ratios. Indeed, s is constant on $\{1, 2, 3\}$ and zero elsewhere; r has singleton support. Their cross term and variance parameter are

$$\langle r, s \rangle_B = \frac{8}{7}, \qquad \kappa(s) = 1 - \operatorname{Var}(Q) = \frac{3}{7}.$$

For completeness, condition on two ones among the four initial bits. At retention parameter κ, each initial one remains one with probability $(1 + \kappa)/2$, and each initial zero becomes one

with probability $(1-\kappa)/2$. The probability of two final ones is $(3+2\kappa^2+3\kappa^4)/8$. Division by $B(2)=3/8$ therefore gives

$$\langle r, T_\kappa r\rangle_B = 1 + \frac{2}{3}\kappa^2 + \kappa^4.$$

At $\kappa = 3/7$ this is

$$\frac{2776}{2401} = \frac{8}{7} + \frac{32}{2401} > \langle r, s\rangle_B.$$

Thus the stronger contraction does not hold for all ultra-log-concave laws. The strong scalar inequality (124) still holds: $2776/2401 < 64/49$. The generating polynomial of Q is $z(2+3z+2z^2)/7$. Its quadratic factor has discriminant $9-16<0$. It is therefore not a Poisson-binomial law, whose generating polynomial is a product of real linear factors.

16.4 Stochastic complements require more than a pivot sign

For a doubly stochastic matrix A and a pivot $a_{ij} < 1$, define

$$B^{ij}_{kl} = a_{kl} + \frac{a_{kj}a_{il}}{1-a_{ij}}, \qquad k \neq i, \quad l \neq j.$$

This stochastic complement is doubly stochastic. For example, its row sum is $1 - a_{kj} + a_{kj}(1-a_{ij})/(1-a_{ij}) = 1$. The column calculation is the same, and all entries are nonnegative. Put

$$R(A) = \frac{\operatorname{per}(A)^2}{\operatorname{per}(AA^{\mathsf{T}})}, \qquad Q_{ij}(A) = \frac{g_n R(B^{ij})}{R(A)}.$$

A universal bound $\min_{a_{ij}<1} Q_{ij}(A) \le 1$ would prove the strong inequality by induction, because $c_n = g_n c_{n-1}$. This bound remains unproved here. The following exact examples exclude selection by positive entries alone, selection by zero entries alone, and two natural averages.

For $n = m+1 \ge 3$, use the symmetric star matrix

$$S_n = \frac{1}{m}\begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & m-1 & 0 & \cdots & 0 \\ 1 & 0 & m-1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & m-1 \end{pmatrix}.$$

A contributing permutation exchanges the center with one leaf and fixes all other leaves. All m contributions have the same weight. Hence

$$\operatorname{per}(S_n) = \frac{(m-1)^{m-1}}{m^m}.$$

The perfect-matching marginal $\pi_{ij} = a_{ij}\operatorname{per}(A_{\widehat{i,j}})/\operatorname{per}(A)$ equals a_{ij}. A center edge occurs in one of the m matchings, and a leaf diagonal occurs in $m-1$ of them. Every positive pivot gives, up to independent row and column permutations,

$$T_m = \frac{1}{m}\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & m-1 & 0 & \cdots & 0 \\ 1 & 0 & m-1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & m-1 \end{pmatrix}$$

of order m. For a leaf diagonal pivot, the update puts $1/m$ in the center diagonal. For a center-edge pivot, the affected leaf row becomes the uniform row. This proves the assertion for both positive-pivot types.

At $n = 6$, exact permanent sums give

	S_6	T_5
per	256/3125	512/3125
per($cdot \cdot^{\mathsf{T}}$)	3222392/48828125	40856/390625
R	40960/402799	32768/127675

and every positive pivot has

$$Q_{ij}(S_6) = \frac{10069975}{9928008} > 1.$$

Thus no positive pivot works, including a pivot of maximum entry. Since the positive entries sum to n and $\pi_{ij} = a_{ij}$, both averages

$$\frac{1}{n}\sum_{ij} a_{ij}Q_{ij}, \qquad \frac{1}{n}\sum_{ij} \pi_{ij}Q_{ij}$$

have this same value, greater than one. A zero pivot between distinct leaves does work: $Q_{12}(S_6) = 53959742125/81341812224 < 1$, with the center indexed by zero. The center zero pivot instead has ratio $4412398635249631/1254003292274688 > 1$. These numbers are finite rational permanent evaluations; the checker computes each sum by subset dynamic programming. Also $R(S_6) > c_6 = 5/324$.

Conversely, for S_3, one has $\mathrm{per}(S_3) = \mathrm{per}(S_3 S_3^{\mathsf{T}}) = 1/4$. All positive pivots have ratio $8/9$, whereas all zero pivots have ratio $200/153 > 1$. For a zero pivot, the complement is a relabeling of $\left(\begin{smallmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{smallmatrix}\right)$. Its permanent ratio is $25/34$; multiplication by $g_3/R(S_3) = 16/9$ gives $200/153$. Thus a useful pivot, if it always exists, must be allowed to come from either support class.

16.5 Averaging minimum norm scales does not replace capacity balancing

For a nonnegative degree-m polynomial in m variables, define

$$\mathcal{C}_F(q) = \inf_{D>0} \frac{N(q(Dx))}{(\det D)^2}.$$

Then $\mathcal{C}_F(q) \le H(q)$. The proposed sufficient deletion bound

$$\sum_j \mathcal{C}_F([x_j]p) \ge n g_n N(p) \tag{182}$$

is false, even for a rank-two doubly stochastic product polynomial. This does not exclude the analogous bound with H.

Let $n \ge 3$, $0 < a < 1$, $b = 1 - a$, and set

$$u = (0, a, b/(n-2), \dots, b/(n-2)), \qquad A_{i1} = u_i, \qquad A_{ij} = \frac{1 - u_i}{n-1} \quad (j \ge 2).$$

This matrix is doubly stochastic and has rank two. With $Y = x_2 + \cdots + x_n$, its product polynomial is

$$p = \sum_{k=0}^{n-1} P_k x_1^k \left(\frac{Y}{n-1}\right)^{n-k}, \qquad P = \mathrm{Bern}(a) + \mathrm{Bin}(n-2, b/(n-2)).$$

Write $B = \mathrm{Bin}(n, 1/n)$ and $C_n(z) = \sum_{k=0}^{n-1}(P_k^2/B_k)z^k$. Expansion of the powers of Y gives $N(p) = c_n C_n(1)$. The special deletion is $q_1 = P_1(Y/(n-1))^{n-1}$, so the identity scale gives

$$\frac{\mathcal{C}_F(q_1)}{g_n N(p)} \le \frac{P_1^2}{g_n^2 C_n(1)}. \tag{183}$$

For an ordinary column $j \ge 2$, put $Z = \sum_{i\ge 2,\, i\ne j} x_i$. Then

$$q_j = \sum_{k=0}^{n-1} \frac{(n-k)P_k}{(n-1)^{n-k}} x_1^k Z^{n-1-k}.$$

At the explicit trial scale $x_1 \mapsto \sqrt{z}\, x_1$, the squared scaling determinant is z. The multinomial expansion of each power of Z gives

$$\frac{\mathcal{C}_F(q_j)}{g_n N(p)} \le \frac{1}{g_n C_n(1)} \sum_{k=0}^{n-1} \frac{P_k^2}{B_k} \frac{n-k}{n-1} \left(\frac{n-2}{n-1}\right)^{n-1-k} z^{k-1}. \tag{184}$$

No optimizing scale is needed for these upper bounds.

Take $n = 700$, $a = 7/10$, and $z = 9/10$. The following integer formulas give a reproducible finite certificate. Set $M = 10(n-2)$, $L = M - 3$, and

$$C_k = \binom{n-2}{k} 3^k L^{n-2-k}, \quad I_k = 3C_k + 7C_{k-1}, \quad D = 10M^{n-2},$$

where out-of-range coefficients are zero. Thus $P_k = I_k/D$ and $\sum I_k = \sum k I_k = D$. Define

$$T_k = I_k^2 k!(n-k)!, \qquad W = \sum_{k=0}^{n-1} T_k (n-1)^k.$$

Then $N(p) = W/[D^2(n-1)^n]$. The special and ordinary trial norms divided by $N(p)$ are, respectively,

$$S = \frac{I_1^2 (n-1)!(n-1)}{W},$$

$$O = \frac{\displaystyle\sum_{k=0}^{n-1} T_k (n-k)(n-2)^{n-1-k}(n-1)^{2k} 9^k 10^{n-1-k}}{(n-1)^n\, 9\, 10^{n-2}\, W}.$$

Exact rational comparison gives

$$\frac{\sum_j \mathcal{C}_F(q_j)}{n g_n N(p)} \le \frac{S + (n-1)O}{n g_n} < \frac{99996}{100000} < 1. \tag{185}$$

This establishes the finite counterexample to (182). The script checks the mass, mean, and strict rational comparison directly. The decimal value $0.9999594712323396\ldots$ is only a description of the exact ratio.

The same failure occurs in all sufficiently large orders of this family. For fixed a, b, the law P tends to $\mathrm{Bern}(a) + \mathrm{Pois}(b)$. The reference B tends to $\mathrm{Pois}(1)$. The limiting collision series is

$$C_\infty(z) = e^{1-2b+b^2 z}[b^2(1+az)^2 + a^2 z]. \tag{186}$$

To verify it, the limiting mass at $k \ge 1$ is $e^{-b} b^{k-1}(b^2 + ak)/k!$, with the same formula interpreted at $k = 0$. Squaring, dividing by $e^{-1}/k!$, and using the exponential series and its first two

derivatives gives (186). The ordinary trial ratio tends to $C_\infty(z)/(zC_\infty(1))$. The special ratio is bounded by e^2, since $P_1 \le 1$, $C_n(1) \ge 1$, and $g_n > e^{-1}$. Its contribution to the average therefore tends to zero. At $a = 7/10$ and $z = 9/10$, the limiting average is

$$e^{-9/1000}\frac{75569}{75010} < 1.$$

Indeed, $75569/75010 = 1 + 559/75010 < 1 + 9/1000 < e^{9/1000}$.

Here are bounds that justify passage to the collision sums. For $k \ge 1$,

$$P_{n,k} \le \frac{b^{k-1}(b^2 + ak)}{k!}, \qquad B_{n,k}^{-1} \le e^{k+2}k!.$$

The first follows by deleting factors at most one from each binomial mass. For the second, use $\binom{n}{k} \ge (n/k)^k$, $k^k/k! \le e^k$, and $(1 - 1/n)^{-n} \le e^2$ for $n \ge 3$. The $k = 0$ terms are bounded separately. The remaining factor in (184), after division by g_n and omission of z^{k-1}, is at most $3e/2$. These bounds give a summable factorial majorant at the fixed positive trial scale. Dominated convergence applies. The failure concerns the average minimum norm only. It leaves open both the maximum-column minimum-norm bound and the capacity-balanced deletion bounds.

16.6 Admissible attenuation need not improve the balanced norm

Proposition 14.13 proves that attenuation of the highest coordinate layer preserves Lorentzianity. The balanced norm is nevertheless neither monotone nor convex along all such paths.

For monotonicity, take $B = \mathrm{Bin}(3, 1/3) = (8, 12, 6, 1)/27$ and attenuate its last likelihood coefficient, giving $(1, 1, 1, \lambda)$. The positive balancing tilt t has mean one exactly when

$$6t^2 + 2\lambda t^3 = 8.$$

Normalize the central likelihood to one. Its squared likelihood norm is

$$\Phi(\lambda) = \frac{8}{27t^2} + \frac{12}{27} + \frac{6t^2}{27} + \frac{\lambda^2 t^4}{27}.$$

At zero, implicit differentiation gives

$$t = \frac{2}{\sqrt{3}}, \qquad t'(0) = -\frac{2}{9}, \qquad \Phi(0) = \frac{26}{27}, \qquad \Phi'(0) = -\frac{4}{81\sqrt{3}} < 0.$$

At $\lambda = 1$, one has $t = 1$ and $\Phi(1) = 1$. Thus filling this coefficient first decreases the balanced norm. This excludes continuous coefficient-filling monotonicity. It does not exclude discrete inclusion monotonicity for supports.

For convexity, use the three-variable polynomial

$$p_\lambda(x, y, z) = \frac{y^3 + 3y^2z + 3yz^2}{6} + \frac{x(y + z)^2}{2} + \lambda\frac{x^2(y + z)}{2}, \qquad 0 \le \lambda \le 1.$$

At one, this is the exponential-generating polynomial on the cubic compositions other than $(3, 0, 0)$ and $(0, 0, 3)$. The support is M-convex: it is the set of nonnegative integer vectors of sum three subject to $\alpha_x, \alpha_z \le 2$; a transfer from an excess coordinate to a deficient coordinate preserves these upper bounds. The polynomial is therefore Lorentzian, and attenuation preserves this property for all λ. Its xyz coefficient is one. Let $G(\lambda) = c_3 N(\widehat{p}_\lambda)$, where the

balanced scale preserves this central coefficient. At zero, use the balanced limit $x(y+z)^2/2$, which gives $G(0) = 4/9$.

Set $y = 1$, $z = t$, $x = u$, and

$$R(t) = \frac{1+3t+3t^2}{6}, \qquad Q(t) = \frac{(1+t)^2}{2}, \qquad L(t) = \frac{1+t}{2}.$$

The capacity objective is $R(t)/(ut) + Q(t)/t + \lambda u L(t)/t$. For $\lambda > 0$, minimization in u gives $u^2 = R/(\lambda L)$. With $\varepsilon = \sqrt{\lambda}$, the remaining objective is

$$\frac{Q(t)}{t} + 2\varepsilon \frac{\sqrt{R(t)L(t)}}{t}.$$

At $\varepsilon = 0$, its unique minimum is $t = 1$, with $(Q/t)''(1) = 1$. The implicit function theorem gives

$$t = 1 + k\varepsilon + O(\varepsilon^2), \qquad k = \frac{3}{14}\sqrt{\frac{7}{6}}, \qquad k^2 = \frac{3}{56},$$

since $[2\sqrt{RL}/t]'(1) = -(3/14)\sqrt{7/6}$. This stationary point is the capacity minimum by convexity in logarithmic scale coordinates. After dividing the scaled polynomial by ut, direct coefficient expansion gives

$$N = 1 + \frac{t^2 + t^{-2}}{2} + \lambda H(t),$$

where

$$H(t) = \frac{L(t)}{R(t)}\left(\frac{1}{6t^2} + \frac{1}{2} + \frac{t^2}{2}\right) + \frac{R(t)}{2L(t)}(t^{-2}+1), \qquad H(1) = \frac{13}{6}.$$

Therefore

$$G'(0+) = \frac{2}{9}\left(2k^2 + \frac{13}{6}\right) = \frac{191}{378}.$$

At $\lambda = 1$, symmetry makes the scales for x, z equal. Their ratio t to the y scale solves $3t^3 - 3t - 1 = 0$, with $9/8 < t < 8/7$. The endpoint norm is

$$G(1) = \frac{2}{9}\left(t^2 + 2 + t^{-2} + \frac{1}{6t^4}\right).$$

The rational enclosure gives

$$t^2 + 2 + t^{-2} + \frac{1}{6t^4} < 2 + \frac{64}{49} + \frac{64}{81} + \frac{2048}{19683} < \frac{359}{84}.$$

Thus $G(1) < 359/378 = G(0) + G'(0+)$. For a convex function, the right derivative at zero cannot exceed the endpoint chord slope. This strict reverse inequality proves failure of convexity.

Finally, support alone cannot give the required upper bound for a product polynomial. Let $C = I + P_4$, where P_4 is a four-cycle permutation matrix. Its product polynomial has central coefficient two and squared Fischer norm 36, so $N(p_C)/(p_C)^2_{\mathbf{1}} = 9$. On the same support S, the exponential-generating polynomial $p_S = \sum_{\alpha \in S} x^\alpha/\alpha!$ is balanced by cyclic symmetry. Its support consists of the center, twelve patterns with $\alpha! = 2$, and two with $\alpha! = 4$. Hence

$$N(p_S) = p_S(\mathbf{1}) = 1 + \frac{12}{2} + \frac{2}{4} = \frac{15}{2} < 9.$$

Both polynomials satisfy the norm inequality, but the proposed comparison between polynomials with the same support has the wrong direction. Extending the support inequality (173) to arbitrary matrix-product polynomials therefore requires control of the coefficients.

17 Verification and reproduction

The principal order-five, order-six, all-order permutation-polynomial, all-order four-block, and exceptional-vector theorems, and rank-two orders two through eight, have Lean proofs. The other results use analytic proofs or exact computational certificates. The accompanying repository contains the programs, outputs, and reproduction commands in `papers/mathematics/gyires/research-notes/complete-support`. The table below identifies the formalized statements.

The order-six and order-seven primary and independent checks cover 5,022,480 coefficient positions. The rank-two calculation covers 33,607 whole parameter intervals through order 15. Its certificates contain 3,524,206 coefficient positions. These are complete finite coefficient checks, not numerical samples of matrix entries. The cubic deletion calculation was regenerated in three forms. The full independent calculation covers all 720 sorting orders. The exact configuration identities were also checked by enumerating every labeled stub matching in seven small parameter cases. Those small cases check the implementation; the counting proof establishes the general identities.

The analytic, scalar, closure, and local polynomial arguments are proved in the text. Their computations are identified at the claims that use them. An exact counterexample only refutes its stated auxiliary assertion. It does not establish the unrestricted target inequality.

17.1 Formal rank-two theorem in orders two through eight

The separate package `papers/mathematics/gyires/formalization/rank-two` proves the conclusions of Theorem 8.1 for every order $2 \le n \le 8$. Its final declaration is `MathAndTheory.Gyires.RankTwo.rank_two_complete_through_eight` in `src/RangeEightComplete.lean`. The hypotheses use actual matrix rank and doubly stochastic membership. No parameter factorization or coefficient-sign hypothesis remains in the final theorem. Zero entries are permitted.

The proof derives the parameter representation, applies separate convexity, and checks every resulting profile interval in Lean's kernel. It proves both Gram contractions, the square contraction, the nonnegative Gyires margin, the normalized left Gram bound, the combined inequality, and equality exactly at the uniform matrix. Its source manifest records the complete source closure, hashes, and import graph. The command `python3 verify.py`, run inside the package after the pinned dependency cache is available, checks every source and the final axiom report. The permitted axioms are `propext`, `Classical.choice`, and `Quot.sound`. The formal range ends at order eight; the written result in orders nine through fifteen still requires complete Lean certificates.

17.2 Order-five coefficient checks

The order-five exact checker reconstructs the permanent from its permutation definition. It checks all 120 weight orderings and the unique reduction to six affine representatives. It verifies the polynomial identities by inverse binomial transforms. Each ordering has 1001 coefficient positions, with 996 positive coefficients and five zeros. The least positive coefficient is 484. The four pure-gap coefficients agree with Lemma 3.9. The computation therefore covers 120,120 positions. The checking program is `papers/mathematics/gyires/verification/check_coefficients.py`.

The Python computation and the Lean proof give separate checks of the coefficient identities.

Neither proof depends on random tests. The repository also preserves a PARI/GP program from [9]; its full set of checks was not rerun.

17.3 Formal certificate for the order-six theorem

The Lean certificate applies to every nonnegative real vector $w\colon \mathrm{Fin}(6) \to \mathbb{R}$ and the matrix $C_{ij} = w(j-i)$, with indices modulo six. It includes zero weights and repeated weights. The declarations use the namespace `MathAndTheory.Gyires.CirculantSix`.

In `SixPrincipal.lean`, `sorted_margin` proves

$$324\,\mathrm{per}(C)^2 - 5\,\mathrm{per}(CC^{\mathsf{T}}) \geq 296\sum_{j=1}^{5}(w_{(j+1)} - w_{(j)})^{12},$$

where $w_{(1)} \leq \cdots \leq w_{(6)}$ is any sorted ordering of the weights. The declaration `sorted_gyires_margin` proves the corresponding combined deficit bound

$$1296\,\mathrm{per}(C)^2 - 5[\mathrm{per}(CC^{\mathsf{T}}) + \mathrm{per}(C^{\mathsf{T}}C) + 2\,\mathrm{per}(C^2)] \geq 1184\sum_{j=1}^{5}(w_{(j+1)} - w_{(j)})^{12}.$$

The declarations `strong_inequality` and `gyires` give the two inequalities without the gap terms. The declarations `strong_equality` and `gyires_equality` prove that either of those inequalities has equality exactly when all six weights are equal. The declaration `gyires_doubly_stochastic` gives the normalized inequality and identifies its equality matrix as the constant matrix with entries $1/6$.

The coefficient 296 is sharp. In `SixSharpness.lean`, `margin_coefficient_le` uses the weights $(0,0,0,0,1,1)$ to prove that this coefficient cannot increase. In `SixNormalizedSharpness.lean`, `normalized_margin_coefficient_le` proves the same statement within the doubly stochastic class, using $(0,0,0,0,1/2,1/2)$.

The source package is `papers/mathematics/gyires/formalization/circulant-six`. This directory contains 3,091 Lean source modules, a relative-path source manifest, and the pinned Lean 4.35.0 and package configuration. From that directory, run `lake exe cache get`, then `python3 verify.py`. The verifier rebuilds the local source set and checks its file hashes. Success requires a zero exit code and `complete: true` in `.verification/result.json`. The root source, `SixTheoremChecks.lean`, restates five matrix conclusions and reports the axioms of ten named theorems. The reported axioms are `propext`, `Classical.choice`, and `Quot.sound`. The package's `THEOREMS.md` maps the results to their source files.

17.4 Permutation-polynomial Lean correspondence

Theorem 5.1 has a complete Lean proof in every dimension. The proof includes the actual cyclic permanent, both Gram products, the transfer identities, and the positive-root lower bound. Explicit formulas cover orders zero through two. Polynomial identities cover orders three through twenty-two, and the weighted trace bound covers every order at least twenty-three. Cycle decomposition then proves the result for every finite permutation. The final theorem has no assumed cycle inequality.

The source package is `papers/mathematics/gyires/formalization/permutation-polynomial`. It contains 405 Lean modules and pins Lean 4.35.0 and Mathlib revision `56a005855ea17e0b0958f73ec801ae140b0a1679`. It is separate from the shared Comparator

selection. Its README gives the dependency, source-integrity, build, and axiom-check commands.

The following declarations use the namespace `MathAndTheory.Gyires.PermutationComplete`. Their source is `src/PermutationComplete.lean` within that package.

Statement	Lean declaration
Both Gram bounds for every cycle order	`cycle_all_gram_bounds`
Both Gram bounds for every permutation	`permutation_polynomial_gram_bounds`
The bound for the original square A^2	`permutation_polynomial_square_bound`
The combined Gyires inequality	`permutation_polynomial_gyires_bound`
Strict square bound for a cycle, $n \geq 23$	`cycle_strict_square_bound`
Strict combined bound for a cycle, $n \geq 23$	`cycle_strict_gyires_bound`

The last two statements require at least one positive weight. The four expanded statements in `src/CompleteAudit.lean` write $A = aI + bQ + cQ^2$, the Mathlib matrix products and permanent, and $n!/n^n$ directly. The formal statements include zero weights, fixed points, and the empty dimension.

The finite proof uses three reflection-reduced gap regions for each order from three through twenty-two. This gives 60 closed regions and 25,050 possible homogeneous coefficient positions. The exact polynomial identities connect their nonnegative coefficients to the actual matrix permanents. Lean checks these identities in its kernel. No coefficient sign is an assumption, and numerical samples are not used as a proof.

The weighted trace argument proves the exact bound at order 23 and every later order by induction. It includes both endpoints of the normalized parameter interval and the zero-weight matrix. The permanent comparison for A^2 uses the original matrix indexing. The ten final axiom reports contain only `propext`, `Classical.choice`, and `Quot.sound`. The proof uses no `native_decide`.

The accompanying ordinary certificate record extends through order 47. It contains 135 reflection-reduced regions and 224,550 coefficient positions. Every coefficient is nonnegative, and every position is positive in orders four through forty-seven. The program `independent_finite.mjs` in `papers/mathematics/gyires/research-notes/extensions/certificates` reproduces that record with argument `47`. The smaller formal certificate range and this wider ordinary computation are separate checks of the same deficit polynomials. The programs `circ_small_check.py` and `sharp_bound_check.py` check the small-order transfer identities and the additional spectral estimate from order 48. Those computational records remain supporting material for the stated ordinary arguments.

17.5 The four-block argument

Theorem 6.1 uses no finite numerical certificate. Its proof derives the likelihood bound for all admissible Bernoulli parameters. The subsequent coefficient identities hold for arbitrary dimension. The equality argument and the closed parameter interval include every boundary case in the stated family. Finite examples can test an implementation, but they do not replace this argument.

17.6 Four-block Lean correspondence

Theorem 6.1 has a complete Lean proof. The formal statement uses Mathlib's matrix product, transpose, permanent, and doubly stochastic predicate. Its hypotheses count the distinct row vectors and column vectors. The proof permits unequal class sizes and zero entries. It also covers one row type, one column type, and finite orders zero and one. The four principal statements are in the namespace `MathAndTheory.GyiresFourBlock`: `contractions`, `margin`, `gyires`, and `gyires_eq_iff`. The independent specifications are in `lean/Challenges/Mathematics/Gyires/CirculantFive.lean`. The matching proofs are in `lean/Solutions/Mathematics/Gyires/FourBlock.lean`.

The supporting library is `lean/MathAndTheory/Mathematics/Gyires/FourBlock.lean`. The probability proof establishes Lemmas 6.2 and 6.3 and the bounds and equality case in Proposition 6.4. It maximizes a linear likelihood difference at a fixed mean. A shifted-binomial coefficient recurrence gives the central bound. The matrix proof then uses a permutation average to prove the contractions. This route proves the full matrix statement without formalizing each intermediate coefficient identity in the ordinary proof.

17.7 Exceptional-vector Lean correspondence

The result for one exceptional column or row has a complete Lean proof for every $n \geq 2$. The predicate states directly that all columns except one are identical, or that all rows except one are identical. The exceptional vector can have arbitrarily many distinct entries, and zero entries are permitted. The formal theorem proves both strong Gram bounds in (61) and Gyires's inequality in (62). Equality in each Gram bound, and in Gyires's inequality, holds exactly at J_n. The matrix square is the original product A^2.

The four statements use the namespace `MathAndTheory.GyiresExceptional`: `strong`, `gram_eq_iff`, `gyires`, and `gyires_eq_iff`. Their independent specifications are in `lean/Challenges/Mathematics/Gyires/CirculantFive.lean`. Their proofs are in `lean/Solutions/Mathematics/Gyires/ExceptionalVector.lean`. The supporting library ends in `MathAndTheory.Mathematics.Gyires.ExceptionalVector`. It proves the one-success probability bounds in both second-moment ranges. It includes deterministic parameters and the exact coefficient identities for the matrix permanents. The four-block theorem covers orders zero and one separately.

17.8 Order-five Lean correspondence

Lean [6] is a *proof assistant*. It checks a formal proof against its statement using basic logical rules. The principal order-five results in Section 3.2 are proved with Lean 4.35.0 and Mathlib [16], its library of formalized mathematics.

The collection library is in `lean/MathAndTheory/Mathematics/GyiresCirculant5`. Its namespace is `MathAndTheory.GyiresCirculant5`. The formalization is based on [9].

In the library, `circ w` is the circulant with entries `w (j - i)`. The definition `P5` is the polynomial in Lemma 3.2. The definitions `gram` and `conv` give the vectors g and h. The type `Fin 5`, with arithmetic modulo five, represents $\mathbb{Z}/5\mathbb{Z}$.

The principal statements are restated, using only Mathlib's definitions, in a *Challenge* file, which states the theorems with their proofs omitted, the file `lean/Challenges/GyiresCirculant5.lean` (namespace `MathAndTheory.GyiresCirculant5`). There a circulant is a matrix of the form `Matrix.circulant v`, whose entries are `v (i - j)`; this is

the transpose of $C(v)$, which describes the same set of matrices. Doubly stochastic matrices are Mathlib's `doublyStochastic`, the permanent is Mathlib's `Matrix.permanent`, $5!/5^5$ is the real quotient `(Nat.factorial 5 : ℝ) / 5 ^ 5`, and J_5 is the constant matrix with entries 5^{-1}. The *Solution* file `lean/Solutions/GyiresCirculant5.lean` proves the same theorems from the library. Lean's Comparator checks that the Solution theorems match the Challenge statements. The verification directory records the statement comparison and the kernel checks by Lean, NanoDa, and con-ron, each with its source revision. The table lists the formalized order-five statements and their supporting declarations. All names use the namespace `MathAndTheory.GyiresCirculant5`.

Statement in this paper	Lean declaration(s)
Theorem 3.4	`strong_inequality`, `strong_inequality_eq_iff`
Theorem 3.5	`strong_margin`, `strong_margin_attained`, `strong_margin_sharp`
Corollary 3.6, homogeneous form	`gyires_homogeneous`, `gyires_homogeneous_eq_iff`
Corollary 3.6, ratio form	`gyires_ratio`, `gyires_ratio_eq_iff`, `gyires_denominator_pos`
Lemma 3.1	`circ_mul_transpose`, `transpose_mul_circ`, `circ_mul_self`
Lemma 3.2	`permanent_circ`
Lemma 2.3, Corollary 3.3	`PermanentGram.permanent_gram_add_ge`, `conv_permanent_le_gram`
Lemma 3.7	`P5_affine`, `gram_affine`, `conv_affine`, `strongDefect_affine`, `gyiresDefect_affine`
Lemma 3.8	`order_affine_cover`, `order_affine_cover_unique`, `canonical_chamber_of_monotone`, `exists_canonical_chamber`
Lemma 3.9, nonnegativity of R_κ	`strong_chamber_0`, ..., `strong_chamber_5`, `strong_chambers`
Theorem 3.5, the example	`strongDefect_twoSupport`, `strongDefect_margin_attained`, `strongDefect_margin_sharp`
Corollary 3.6, positive denominator	`positiveDenominator`
Section 18, the two-point example	`P5_twoSupport`, `P5_gram_twoSupport`, `P5_conv_twoSupport`, `gyiresDefect_twoSupport`, `gyiresDenominator_twoSupport`, `permanent_twoSupport_half`, `permanent_gram_twoSupport_half`, `permanent_square_twoSupport_half`, `normalizedGyiresRatio_twoSupport_half`

Lean's `#print axioms` command lists the logical assumptions used by each theorem. The recorded checks identify only three standard Lean axioms (`propext`, `Classical.choice`, `Quot.sound`); no step is left unproved, and the proofs use neither `native_decide` nor any other evaluation by compiled code. The formal proof of Lemma 3.9 follows the organization described after it: the six chamber polynomials are expanded in z_0 and z_1, and each of the resulting polynomials in z_2, z_3, z_4 is written with explicit nonnegative integer coefficients, proved equal to the corresponding expansion coefficient by the `ring` tactic and nonnegative by the `positivity` tactic. These explicit coefficients were produced by a computer program, but Lean does not rely on that program: it checks every identity and every inequality itself. The formal proof of Lemma 2.3 is not the occupancy argument given here but a sum-of-squares argument over permutations, adapted from the formalization of Ozols [15].

Several supporting results are also formalized: the case $n = 2$ discussed in Section 18 (`gyires_two`); the characterization of the doubly stochastic circulants of order five and their description as the combinations $\sum_k w_k Q^k$ (`circulant_mem_doublyStochastic_iff`, `circulant_eq_sum_pow`); and, for real square matrices of every order, the reduction of Gyires's inequality to the strong form in Section 2 (`gyires_of_strong`).

17.9 General supporting Lean results

The modules in `lean/MathAndTheory/Mathematics/Gyires` prove several comparisons used beyond order five. The module `PermanentCauchySchwarz.lean` proves Lemma 2.2 for real square matrices of every finite size. Its proof uses the nonnegative permutation bilinear form and a quadratic discriminant.

The module `PermanentPermutation.lean` proves that independent row and column permutations preserve both strong Gram bounds. It then applies those bounds to the square of the permuted matrix. It does not assert invariance of $\mathrm{per}(A^2)$ under independent row and column permutations.

The module `DimensionConstant.lean` proves the comparison of the constants c_n for positive block sizes. The comparison is strict for at least two blocks. This supplies the constant comparison in the cycle and direct-sum arguments. The module `BlockPermanent.lean` proves Lemma 9.1. It proves the permanent product identity and both strong bounds for finite block sums. The block sizes can differ, and independent row and column permutations are permitted. The bounds are strict for at least two positive-dimensional doubly stochastic blocks. The proof also gives Gyires's inequality for the square of the resulting matrix. These closure results do not prove the required block hypotheses for either all-order structured class.

17.10 Limits of the formal verification

The formal proof establishes $R_\kappa(z) \geq 0$ for all $z \geq 0$, which is the part of Lemma 3.9 that the theorems need. The finer statements of that lemma (the numbers 996 and 5 of positive and zero coefficients, the positions of the zeros, the four pure-power coefficients, and the least positive coefficient 484) are not stated in Lean; the exact computation above verifies them. The broader matrix and polynomial results are not fully formalized. The Cauchy–Schwarz proof in Section 2 uses the coefficient inner product. The Lean proof uses the permutation bilinear form and a quadratic discriminant. The formal verification therefore applies to the listed declarations, rather than to every argument in this paper.

18 Comparison of the results and remaining questions

The common reduction in Section 2 explains how the matrix arguments address the same conjecture. The circulant and permutation-polynomial proofs establish a strong Gram-permanent bound. The four-block and finite rank-two proofs establish the sharper contractions $\mathrm{per}(AA^{\mathsf{T}}) \leq \mathrm{per}(A)$ and $\mathrm{per}(A^{\mathsf{T}}A) \leq \mathrm{per}(A)$ on their respective classes. Those contractions are not valid for all matrices covered by the circulant results.

18.1 A two-point example and a distinction between the bounds

For the order-five weight vector $w = (t, t, 0, 0, 0)$,

$$g(w) = (2t^2, t^2, 0, 0, t^2), \qquad h(w) = (t^2, 2t^2, t^2, 0, 0),$$

and

$$P(w) = 2t^5, \qquad P(g(w)) = P(h(w)) = 84t^{10}, \qquad D_s(w) = D_G(w) = 484t^{10}.$$

This is the sharpness example in Theorem 3.5. At $t = 1/2$, the matrix is doubly stochastic and

$$\operatorname{per}(C) = \frac{1}{16}, \qquad \operatorname{per}(CC^{\mathsf{T}}) = \operatorname{per}(C^2) = \frac{21}{256}.$$

Thus $\operatorname{per}(CC^{\mathsf{T}}) > \operatorname{per}(C)$, although the strong bound and Gyires's inequality both hold. The Gyires ratio is $1/21 > 24/625$. The two-point example also belongs to the permutation-polynomial family. It shows why the four-block contraction theorem must retain its row-type and column-type hypotheses. An earlier order-four contraction counterexample is attributed to Newman in the account of Best and Wanless [1].

18.2 The elementary order-two case

For $A = \left(\begin{smallmatrix} a & 1-a \\ 1-a & a \end{smallmatrix}\right)$ with $0 \le a \le 1$, put $s = a^2 + (1-a)^2$ and $t = 2a(1-a)$. Then $\operatorname{per}(A) = s$, and each permanent of a product in Gyires's denominator is $s^2 + t^2$. The ratio inequality reduces to $s \ge t$, which is $(2a-1)^2 \ge 0$. Equality holds only at $a = 1/2$. This elementary case is also covered by the four-block theorem.

18.3 What remains outside the proofs

For a general nonnegative circulant, normality gives $CC^{\mathsf{T}} = C^{\mathsf{T}}C$. The permanent comparison still gives $\operatorname{per}(C^2) \le \operatorname{per}(CC^{\mathsf{T}})$. The exact certificates for unrestricted circulant weights cover orders five through seven. The strong Gram bound for unrestricted circulant weights in arbitrary order remains unproved here. The three-term transfer construction does not supply that step.

A general matrix of rank two can have more than two row types and more than two column types. The four-block theorem does not cover that full class. The finite rank-two theorem covers orders through fifteen. The all-order rank-two assertion remains unproved here. The unrestricted Gyires conjecture and the general strong Gram inequality remain outside this manuscript.

The general order-seven circulant result and rank-two orders 9 through 15 still require complete Lean proofs. Each proof must connect its polynomial identities and coefficient bounds to the actual matrix theorem. The permutation-polynomial proof includes this connection in every dimension.

19 Research process and AI disclosure

This work is fully autonomous agentic research. The initial research and proofs were generated autonomously by GPT 6 with Pro Reasoning in ChatGPT. GPT 6 Astra in Codex autonomously produced the formalizations and corrections. Earlier source records also identify the order-five formalization as work done in Claude Code. GPT 6.1 Sol in Codex generated the order-eight

rank-two certificate modules and matrix extension on 29 October 2026. The author checked the proofs to the best of their mathematical abilities; this check is not expert validation or peer review. Dev D. Goyal is responsible for publication. This preprint has not been externally peer reviewed. No priority claim is made for the broader results.

Opus 5.5 by Anthropic used with Claude Code made the editorial revision of 1 November 2026 to the front matter, the abstract and this section; it did not change the mathematics. Section 17 states the formal scope and the reproduction commands, and Section 18.3 states what remains outside the proofs.

Acknowledgements

The conjecture is due to B. Gyires. The order-five formalization of the permanent comparison adapts Chris Ozols's proof [15]. Its Apache License 2.0 notices are preserved in the source files. The formal proofs use Lean and Mathlib.

A Additional results for permanent and polynomial inequalities

This appendix gives exact counterexamples and supporting reductions related to Gyires's inequality. Each counterexample states the assertion that fails. None is a counterexample to Gyires's inequality for nonnegative doubly stochastic matrices. The results below have not been formalized in Lean.

A.1 Definitions

Let $c_n = n!/n^n$, and let J_n have every entry equal to $1/n$. A doubly stochastic matrix has nonnegative entries, and every row and column sum is one. For a real square matrix A, put

$$p_A(x) = \prod_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_j = \sum_{|\alpha|=n} p_\alpha x^\alpha, \qquad N(p) = \sum_{\alpha} \alpha! p_\alpha^2, \qquad \alpha! = \prod_j \alpha_j!.$$

The permanent coefficient identity gives

$$[x_1 \cdots x_n] p_A = \operatorname{per}(A), \qquad \langle p_A, p_B \rangle_F = \sum_{\alpha} \alpha! p_\alpha q_\alpha = \operatorname{per}(AB^{\mathsf{T}}).$$

Here q_α are the coefficients of p_B. The squared Fischer norm is $N(p) = \langle p, p \rangle_F$. Thus the strong Gram inequality is

$$\operatorname{per}(A)^2 \geq c_n N(p_A) = c_n \operatorname{per}(AA^{\mathsf{T}}). \tag{187}$$

For the original symmetrized denominator, write

$$Q(A) = \frac{\operatorname{per}(AA^{\mathsf{T}}) + \operatorname{per}(A^{\mathsf{T}}A) + 2\operatorname{per}(A^2)}{4}.$$

Gyires's inequality is $\operatorname{per}(A)^2 \geq c_n Q(A)$. Applying (187) to both A and A^{T} implies this inequality by permanent Cauchy–Schwarz.

A degree-m polynomial q in m variables is balanced when $\partial_j q(\mathbf{1}) = q(\mathbf{1})$ for every j. For a polynomial with nonnegative coefficients, its capacity is

$$\operatorname{cap}(q) = \inf_{x_j>0} \frac{q(x)}{x_1 \cdots x_m}.$$

At a positive diagonal capacity minimizer D, the polynomial $q(Dx)/\det D$ is balanced. Write $H(q) = N(q(Dx))/(\det D)^2$ for its squared Fischer norm. When a finite minimizer does not exist, use the balanced polynomial limit under determinant-one capacity-minimizing scales. The examples that use H below give explicit balancing formulas. The notation $[x_j]p$ extracts exactly the first power of x_j, leaving a polynomial in the other variables.

A polynomial is real stable if it has no zero when every variable has positive imaginary part. A nonnegative homogeneous stable polynomial is Lorentzian. In particular, each quadratic derivative of a degree-n Lorentzian polynomial has a Hessian with at most one positive eigenvalue. The examples below use this necessary condition. When that condition is used to establish Lorentzianity, the support condition is also stated. For a set $S \subseteq \mathbb{N}^m$ of fixed coordinate sum, M-convexity means the following exchange condition. For all $\alpha, \beta \in S$ and every i with $\alpha_i > \beta_i$, there is a j with $\alpha_j < \beta_j$ such that both $\alpha - e_i + e_j$ and $\beta + e_i - e_j$ belong to S. Here e_i is the i-th unit vector.

A.2 The central coefficient need not dominate repeated columns

Consider

$$A = \frac{1}{4}\begin{pmatrix} 1 & 1 & 2 \\ 0 & 2 & 2 \\ 3 & 1 & 0 \end{pmatrix}.$$

This matrix is doubly stochastic. Direct permanent sums give

$$\operatorname{per}(A) = \frac{5}{16}, \qquad \operatorname{per}(A_{[:,(1,3,3)]}) = \frac{3}{8} > \frac{5}{16}.$$

The matrix $A_{[:,(1,3,3)]}$ repeats column three and omits column two. For a general multiplicity vector α, the repeated-column permanent equals $\alpha! p_\alpha$. Therefore $\alpha! p_\alpha \le p_{\mathbf{1}}$ is false. Summing this pointwise comparison cannot prove (187). The actual strong deficit is positive:

$$\operatorname{per}(AA^{\mathsf{T}}) = \frac{17}{64}, \qquad \operatorname{per}(A)^2 - c_3 \operatorname{per}(AA^{\mathsf{T}}) = \frac{89}{2304}.$$

A.3 Pairwise averaging need not decrease the strong ratio

Write $R(A) = \operatorname{per}(A)^2/\operatorname{per}(AA^{\mathsf{T}})$. For

$$A = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \qquad B = \frac{1}{4}\begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix},$$

the matrix B averages rows one and three of A. Both matrices are doubly stochastic, and

$$\operatorname{per}(A) = \operatorname{per}(B) = \frac{1}{4}, \quad \operatorname{per}(AA^{\mathsf{T}}) = \frac{1}{4}, \quad \operatorname{per}(BB^{\mathsf{T}}) = \frac{15}{64}.$$

Hence $R(A) = 1/4 < 4/15 = R(B)$.

A separate column example is

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

The second matrix averages columns one and two. Exact evaluation gives

$$R(C) = \frac{1089}{3841} < \frac{169}{590} = R(\widetilde{C}).$$

Thus arbitrary row or column averaging cannot be assumed to decrease R toward $R(J_n) = c_n$. These examples do not exclude a selected sequence of averaging steps.

A.4 Normalized coefficients are not multivariate Jensen-concave

Let

$$A = \frac{1}{5}\begin{pmatrix} 2 & 0 & 3 \\ 0 & 3 & 2 \\ 3 & 2 & 0 \end{pmatrix}, \qquad 125p_A = (2x + 3z)(3y + 2z)(3x + 2y).$$

Put $D_\alpha = \alpha! p_\alpha$. Expansion gives

$$D_{111} = \frac{35}{125}, \qquad D_{210} = D_{021} = D_{102} = \frac{36}{125}.$$

The average of the three latter exponent vectors is $(1, 1, 1)$. Nevertheless,

$$\log D_{111} < \frac{1}{3}(\log D_{210} + \log D_{021} + \log D_{102}).$$

The polynomial is a balanced stable product of nonnegative linear forms. These properties do not give full multivariate Jensen concavity. The three weights in this comparison are $1/3$. They are not the actual occupancy probabilities, so this example alone does not refute an entropy inequality with those probabilities.

A.5 The collision polynomial can lose stability and balance

For a product polynomial, define

$$h_A(x) = \sum_\alpha \alpha! p_\alpha^2 x^\alpha = \operatorname{per}(A \operatorname{diag}(x) A^{\mathsf{T}}).$$

Use the matrix C in Section A.3. After multiplication by 5^6, its collision polynomial is

$$h(x,y,z) = 96x^3 + 392x^2y + 392x^2z + 288xy^2 \\ + 1089xyz + 288xz^2 + 648y^2z + 648yz^2.$$

Its mass is 3841. Its three first-moment numerators are $3521, 4001, 4001$. Thus its normalized mean is

$$\left(\frac{3521}{3841}, \frac{4001}{3841}, \frac{4001}{3841}\right) \neq (1,1,1).$$

Also,

$$\operatorname{Hess}(\partial_y h) = \begin{pmatrix} 784 & 576 & 1089 \\ 576 & 0 & 1296 \\ 1089 & 1296 & 1296 \end{pmatrix}$$

has determinant -120932352 and trace 2080. It has two positive eigenvalues and one negative eigenvalue. The quadratic derivative therefore violates the Lorentzian Hessian condition. Differentiation preserves stability, so h is neither stable nor Lorentzian.

The same example also excludes an arbitrary unscaled singleton deletion. Terms with x-exponent one have total collision weight 1665. Therefore

$$\frac{N([x]p_C)}{N(p_C)} = \frac{1665}{3841} < \frac{4}{9}, \qquad 14985 < 15364.$$

This last calculation concerns the unscaled norm. It does not determine the norm after capacity balancing.

For a nonnegative matrix A with $N(p_A) > 0$, the normalized collision law has a useful exact mixture representation. Choose a permutation σ with probability proportional to $\prod_i (AA^{\mathsf{T}})_{i,\sigma(i)}$. Conditional on this permutation, choose labels independently, with row-i probabilities proportional to $a_{ij}a_{\sigma(i),j}$. Group the labels by occupancy. The resulting occupancy has mass $\alpha! p_\alpha^2 / N(p_A)$. Permutations with zero weight can be omitted. The conditional row distributions need not form a doubly stochastic matrix. Thus van der Waerden's theorem does not apply to each component merely from this representation.

A.6 A global exponential trace lower bound fails

The bound

$$\frac{\operatorname{per}(A)}{c_n} \geq \exp\left(\frac{1}{2}\|A - J_n\|_F^2\right) \tag{188}$$

fails for

$$A = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Here $\operatorname{per}(A)/c_3 = 9/8$ and $\|A - J_3\|_F^2 = 1/2$, whereas $e^{1/4} > 5/4 > 9/8$.

More generally, take $A = (nJ_n - I)/(n-1)$. If D_n is the number of derangements of n elements, then

$$\frac{\operatorname{per}(A)}{c_n} = \frac{D_n n^n}{n!(n-1)^n}, \qquad \|A - J_n\|_F^2 = \frac{1}{n-1}.$$

At $n = 5$, the ratio is $6875/6144$, but $e^{1/8} > 9/8 > 6875/6144$. These failures exclude (188), not weaker trace bounds.

A.7 Fourier magnitudes and a fixed transfer gap are insufficient

For a circulant $C(w)_{ij} = w_{j-i \bmod n}$, its Gram matrix depends only on its cyclic autocorrelation. At order three, let

$$w = (4, 1, 1), \qquad v = (3, 3, 0).$$

Both autocorrelations are $(18, 9, 9)$, but

$$\operatorname{per} C(w) = 78, \qquad \operatorname{per} C(v) = 54, \qquad \operatorname{per}(C(w)C(w)^{\mathsf{T}}) = 11664.$$

These values follow from $\operatorname{per} C(a, b, c) = a^3 + b^3 + c^3 + 3abc$. After division by six, $C(w)$ has eigenvalues $1, 1/2, 1/2$. The nontrivial eigenvalues of $C(v)/6$ have magnitudes $1/2$ and phases $\pm\pi/3$. Replacing phases by zero can therefore increase the permanent. A fixed-Gram argument must retain phase information.

A separate limitation follows at J_n:

$$\frac{\operatorname{per}(J_n J_n^{\mathsf{T}})}{\operatorname{per}(J_n)^2} = \frac{n^n}{n!}, \qquad \left(\frac{n^n}{n!}\right)^{1/n} \longrightarrow e.$$

Thus no bound

$$\operatorname{per}(CC^{\mathsf{T}}) \le b_n K^n \operatorname{per}(C)^2, \qquad b_n^{1/n} \longrightarrow 1,$$

can cover every circulant with one fixed $K < e$. A transfer bound valid for arbitrary support must allow this gap to close.

A.8 Two stability failures for circulant coefficient constructions

First take $C = I + P_4$, with a cyclic orientation for which

$$p_C = (x_0 + x_1)(x_1 + x_2)(x_2 + x_3)(x_3 + x_0).$$

For $F_C = \sum \alpha! p_\alpha^2 x^\alpha$, expansion gives

$$\partial_{x_0}^2 F_C = 4(x_1x_2 + x_1x_3 + 2x_2^2 + x_2x_3).$$

At $x_1 = x_3 = 1, x_2 = t$, this is $4(2t^2 + 2t + 1)$, with roots $(-1 \pm i)/2$. Its Hessian block is

$$\begin{pmatrix} 0 & 4 & 4 \\ 4 & 16 & 4 \\ 4 & 4 & 0 \end{pmatrix},$$

with eigenvalues $-4, 10 - 2\sqrt{17}, 10 + 2\sqrt{17}$. Thus factorial-weighted coefficient squaring does not preserve stability or Lorentzianity, even for a circulant product. Cyclic symmetry still gives balance of F_C. In fact, weighted arithmetic–geometric mean gives $F_C(x) \geq F_C(\mathbf{1}) \prod_j x_j$ for $x_j > 0$. Balance does not repair the failed stability condition.

Second, let P be a four-cycle and put

$$R_0 = I + P + P^2 + P^3, \quad R_1 = P + P^2 + P^3, \quad R_2 = P^2 + P^3, \quad R_3 = P^3.$$

Define the mixed gap coefficient

$$f(y) = [z_0 z_1 z_2 z_3] \prod_{i=0}^{3} \left(\sum_{j=0}^{3} \sum_{r=0}^{3} z_r (R_r)_{ij} y_j \right).$$

The Hessian of $\partial_{y_0}^2 f$ is

$$H = \begin{pmatrix} 24 & 36 & 48 & 48 \\ 36 & 48 & 56 & 56 \\ 48 & 56 & 56 & 60 \\ 48 & 56 & 60 & 48 \end{pmatrix}.$$

Its leading 3×3 block has determinant -384 and positive trace. More explicitly, at

$$y_3 = 0, \qquad (y_0, y_1, y_2) = (-30 + t, 38 + t, -13 + t),$$

the quadratic is $204(1 + t^2)$. This has nonreal roots. Differentiation and real specialization would preserve stability or give zero. Hence this extracted polynomial is not stable. The example does not give a negative coefficient in a circulant sorted-gap deficit.

A.9 Transpose averaging does not restore stability or collision balance

Let

$$A = \frac{1}{6} \begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 5 \\ 2 & 4 & 0 \end{pmatrix}, \qquad A_0 = 6A, \qquad U = p_{A_0} + p_{A_0^{\mathrm{T}}}.$$

The Hessian of $\partial_y U$ is

$$H = \begin{pmatrix} 96 & 36 & 168 \\ 36 & 0 & 120 \\ 168 & 120 & 120 \end{pmatrix}, \qquad \det H = -86400, \quad \operatorname{tr} H = 216.$$

It has two positive eigenvalues. An explicit stability obstruction uses $x = (-12, 10, 5)$ and $e = (1, 1, 1)$:

$$x^{\mathsf{T}}Hx = 24, \quad x^{\mathsf{T}}He = 0, \quad e^{\mathsf{T}}He = 864, \qquad \tfrac{1}{2}(x+te)^{\mathsf{T}}H(x+te) = 12 + 432t^2.$$

At $t = i/6$, every coordinate of $x + te$ has positive imaginary part, and this quadratic vanishes. Thus U is neither stable nor Lorentzian.

If $U = \sum r_\alpha x^\alpha$, its collision weights satisfy

$$\sum_\alpha \alpha! r_\alpha^2 = 54000, \qquad \frac{\sum_\alpha \alpha\, \alpha! r_\alpha^2}{54000} = \left(\frac{117}{125}, \frac{129}{125}, \frac{129}{125}\right).$$

Transpose averaging therefore does not restore balanced collision moments. The matrix itself satisfies Gyires with positive deficit

$$4\,\mathrm{per}(A)^2 - 4c_3 Q(A) = \frac{169}{486}.$$

A.10 An operator inequality stronger than the scalar target fails

Let $T = A^{\otimes n}$. Let E project orthogonally onto tensor basis vectors whose indices are all distinct. In the permutation basis, $S = ETE$ has entries

$$S_{\sigma,\tau} = \prod_{i=1}^{n} a_{\sigma(i),\tau(i)}.$$

Every row and column sum is $\mathrm{per}(A)$. If v is the unit uniform vector on this subspace, then $Sv = S^{\mathsf{T}}v = \mathrm{per}(A)v$. For symmetric A, the scalar target would follow from

$$S^2 \succeq c_n E(A^2)^{\otimes n} E. \tag{189}$$

Use the symmetric order-three matrix from Section A.6. Order the basis as

$$(123),\ (132),\ (213),\ (231),\ (312),\ (321).$$

Put $w = e_{123} + e_{132}$ and $G = E(A^2)^{\otimes 3}E$. Exact matrix products give

$$\langle w, S^2 w\rangle = \frac{1}{16}, \qquad \langle w, Gw\rangle = \frac{5}{16}, \qquad \langle w, (S^2 - \tfrac{2}{9}G)w\rangle = -\frac{1}{144}.$$

Thus (189) fails even on a nonnegative vector. On the required uniform vector v, the scalar strong deficit is instead $1/144$. The scalar problem requires information specific to that vector.

A.11 Positive-semidefinite completion does not give coefficientwise positivity

For a doubly stochastic A, set

$$S_A = \begin{pmatrix} 0 & A \\ A^{\mathsf{T}} & 0 \end{pmatrix}, \qquad B_t = \frac{I_{2n} + tS_A}{1+t}, \quad 0 \le t \le 1.$$

The bound $\|A\|_2 \le 1$ makes B_t positive semidefinite and doubly stochastic. Also,

$$\operatorname{per}(I_{2n}+tS_A)=\sum_{k=0}^{n}t^{2k}\sum_{|I|=|J|=k}\operatorname{per}(A[I,J])^2.$$

The identity entries fix the unused rows and columns. The two cross-block matchings produce the same minor permanent twice. Since $0 \preceq B_t^2 \preceq B_t$, permanent monotonicity in the positive-semidefinite order gives

$$\operatorname{per}((I_{2n}+tS_A)^2) \le (1+t)^{2n}\operatorname{per}(I_{2n}+tS_A).$$

This valid inequality on $[0,1]$ does not imply coefficientwise positivity. The coefficient of t^{4n} in its right side minus its left side is

$$\operatorname{per}(A)^2-\operatorname{per}(AA^{\mathsf{T}})\operatorname{per}(A^{\mathsf{T}}A).$$

For $A=(I+P_4)/2$, it equals

$$\frac{1}{64}-\left(\frac{9}{64}\right)^2=-\frac{17}{4096}.$$

The same matrix gives a distinct limit for duplication:

$$\operatorname{per}(A)=\frac{1}{8},\qquad \operatorname{per}(AA^{\mathsf{T}})=\frac{9}{64},\qquad \operatorname{per}(A\otimes \mathbf{1}_{2\times 2})=\frac{9}{8},$$

$$\operatorname{per}(A)^2-\frac{\operatorname{per}(A\otimes \mathbf{1}_{2\times 2})}{\binom{8}{4}}=-\frac{1}{2240}.$$

Thus the rank-two duplication bound cannot extend to all nonnegative doubly stochastic matrices. This example is also stated in Remark 1.3 of Kunisky, Spielman, and Yu, *Inequalities for rank-two permanents and finite free convolutions*, arXiv:2608.28520v1. The strong deficit for this matrix is $5/2048>0$.

A.12 Matrix symmetrization and signed contractions do not give the required reductions

For

$$A=\frac{1}{5}\begin{pmatrix}0&0&5\\1&4&0\\4&1&0\end{pmatrix},\qquad B=\frac{A+A^{\mathsf{T}}}{2},$$

exact evaluation gives

$$\frac{\operatorname{per}(A)^2}{Q(A)}=\frac{289}{321},\qquad \frac{\operatorname{per}(B)^2}{Q(B)}=\frac{110889}{122956}.$$

The first ratio minus the second is $-61085/39468876<0$. Plain symmetrization therefore need not decrease the original Gyires ratio.

The nonnegative-entry hypothesis also cannot be replaced by symmetry, contraction, and the van der Waerden lower bound. The signed matrix

$$C = \frac{1}{6}\begin{pmatrix} -2 & 4 & 4 \\ 4 & 1 & 1 \\ 4 & 1 & 1 \end{pmatrix}$$

is symmetric, has row and column sums one, and has eigenvalues $1, -1, 0$. Yet

$$\operatorname{per}(C) = \frac{5}{18} > c_3, \qquad \operatorname{per}(C^2) = \frac{1}{2}, \qquad \operatorname{per}(C)^2 - c_3 \operatorname{per}(C^2) = -\frac{11}{324}.$$

In the centered rank-two representation $C_{ij} = (1 + x_i y_j)/3$, take $x = (\sqrt{2}, -1/\sqrt{2}, -1/\sqrt{2})$ and $y = -x$. Then the normalized permanent is $5/4$, and the normalized Gram permanent is $9/4$. Centering and a singular-value bound do not replace entrywise nonnegativity.

A.13 The occupancy entropy bridge fails with the actual probabilities

For doubly stochastic A, the coefficients p_α of p_A form a probability law. Put

$$q_\alpha = \frac{c_n}{\alpha!}, \qquad r_\alpha = \frac{p_\alpha}{q_\alpha}, \qquad D(p\|q) = \sum_{p_\alpha > 0} p_\alpha \log r_\alpha.$$

The strong target is $E_q r^2 \le r_{\mathbf{1}}^2$. The proposed auxiliary inequality $D(p\|q) \le \log r_{\mathbf{1}}$ fails for $A = (I + P_n)/2$ and every $n \ge 4$.

To compute its exact gap, use independent fair bits $X_1, \dots, X_n$. The occupancy is $\alpha_j = 1 - X_j + X_{j-1}$, with cyclic indices. Every nonconstant binary word is determined by its cyclic differences. The two constant words both give $\mathbf{1}$. Hence

$$p_{\mathbf{1}} = 2^{1-n}, \qquad p_\alpha = 2^{-n} \quad \text{for every other supported occupancy.}$$

If K counts transitions of either fixed orientation around the cycle, then $\alpha! = 2^K$ and $EK = n/4$. The entropy is

$$-\sum_{p_\alpha > 0} p_\alpha \log p_\alpha = (n - 2^{1-n}) \log 2.$$

Therefore

$$D(p\|q) = -\frac{3n}{4} \log 2 + 2^{1-n} \log 2 - \log c_n,$$

$$D(p\|q) - \log r_{\mathbf{1}} = \left(\frac{n}{4} - 1 + 2^{1-n}\right) \log 2 > 0 \quad (n \ge 4).$$

At $n = 4$, $r_{\mathbf{1}} = 4/3$ and $D(p\|q) = \log(4/3) + (\log 2)/8$. Still, $E_q r^2 = 3/2 < 16/9 = r_{\mathbf{1}}^2$. This failure concerns the entropy bridge. It makes no claim about a separate proposed comparison between order-two Rényi divergence and twice the relative entropy.

A.14 Balance and negative dependence do not replace product structure

Balance alone is insufficient. The law assigning mass $1/n$ to each pure occupancy ne_j, for $n \geq 2$, has mean $\mathbf{1}$ and zero central mass. Its polynomial is $\sum_j x_j^n/n$, with $N = n!/n > 0$. Thus its central coefficient cannot satisfy the norm inequality.

A stronger example also has negative association and bounded moments. For $0 \leq \theta \leq 1$, let $p^{(\theta)}$ put mass θ at $(1,1,1)$ and mass $(1-\theta)/6$ at each permutation of $(2,1,0)$. Its mean is $\mathbf{1}$, and

$$\mathrm{Cov}(p^{(\theta)}) = (1-\theta)(I_3 - J_3) \preceq I_3 - J_3 = \mathrm{Cov}(q).$$

Every mixed falling-factorial moment is at most one. The first moments equal one. At degree two the moments are $2(1-\theta)/3$ for a repeated coordinate and $(2+\theta)/3$ for distinct coordinates. At degree three they are $(1-\theta)/3$ for pattern $(2,1,0)$, θ for $(1,1,1)$, and zero for $(3,0,0)$. Higher moments vanish.

To check negative association, condition on α_1. For values zero and two, the other pair is uniform on $\{(1,2),(2,1)\}$ and $\{(0,1),(1,0)\}$, respectively. At value one, it is a mixture of the uniform law on $\{(0,2),(2,0)\}$ and the point mass at $(1,1)$. The upper conditional law dominates the middle law coordinatewise, and the middle dominates the lower law. This follows by coupling the two unequal points with the corresponding upper or lower points, and coupling $(1,1)$ with an equal mixture of those points. Thus every increasing function of the other coordinates has conditional expectation decreasing in α_1. Symmetry gives the same property for each coordinate. In three coordinates, disjoint nonempty coordinate sets always have a singleton side. This proves negative association.

At $\theta = 3/13$,

$$p_{111} = \frac{3}{13} > c_3, \qquad N(p) = \theta^2 + \frac{(1-\theta)^2}{3} = \frac{127}{507}, \qquad c_3 N(p) - p_{111}^2 = \frac{11}{4563} > 0.$$

This is an abstract occupancy law, not a doubly stochastic product polynomial. For its generating polynomial, the Hessian of the x-derivative has eigenvalue $(1-4\theta)/3 = 1/39$ in the antisymmetric (y,z) direction. Its complementary block has another positive eigenvalue. The polynomial is neither stable nor Lorentzian. The example excludes a replacement of product structure by just the stated dependence and moment conditions.

A.15 A mixed-coefficient estimate fails for Lorentzian cubics

For a balanced ternary cubic, write $b = [xyz]p$, $U_i = [x_i^3]p$, and

$$[x_j]p = a_j x_k^2 + b x_k x_\ell + c_j x_\ell^2.$$

The squared Fischer norm after capacity balancing of this binary quadratic is $H_j = b^2 + 4a_j c_j$. Direct expansion gives

$$G(p) := 3\sum_j H_j - 4N(p) = 5b^2 - \sum_j ((a_j + c_j)^2 + 7(a_j - c_j)^2) - 24\sum_i U_i^2.$$

The separate proposed estimate

$$\sum_j ((a_j + c_j)^2 + 7(a_j - c_j)^2) \le 3b^2 \tag{190}$$

is false. Take

$$p = (y+z)\left[\frac{(y+z)^2}{4\sqrt{3}} + \frac{x(y+z)}{2} + \frac{x^2}{\sqrt{3}}\right].$$

It is balanced, $b = 1$, $U_x = 0$, and $p(\mathbf{1}) = 2 + 4\sqrt{3}/3$. Its y- and z-derivatives are squares of real linear forms. Its x-derivative is $(y+z)((y+z)/2 + 2x/\sqrt{3})$. The support is the cubic simplex with only $(3,0,0)$ removed, which is M-convex. These quadratic derivatives have at most one positive Hessian eigenvalue, so p is Lorentzian.

The exact values are

$$N(p) = \frac{13}{3}, \quad H_x = H_y = H_z = 2, \quad G(p) = \frac{2}{3}, \quad \sum_j ((a_j + c_j)^2 + 7(a_j - c_j)^2) = \frac{10}{3} > 3.$$

The desired average-deletion inequality still holds. The quadratic factor in variables $x, s = y+z$ has discriminant $-1/12$. Therefore p is not real stable. This example does not refute the stable-only form of (190).

A.16 Three limits of scalar likelihood reductions

Let P_x be the sum of independent Bernoulli variables with parameters $x \in [0,1]^n$. Fix $\sum_i x_i = t$, with $0 < t < n$, and let $B = \mathrm{Bin}(n, t/n)$. A likelihood ratio is $r(k) = P_x(k)/B(k)$.

Noninteger interpolation. Take $n = 3, t = 7/4$ and $P = \mathrm{Bin}(2, 7/8)$, with one additional zero Bernoulli parameter. Relative to $B = \mathrm{Bin}(3, 7/12)$,

$$r = (27/125, 18/25, 9/5, 0), \qquad E_P r = \frac{1539}{1000} > \frac{1}{4}r(1) + \frac{3}{4}r(2) = \frac{153}{100}.$$

Thus linear interpolation at a noninteger mean does not preserve the integer-mean collision-norm bound.

Relative log-concavity after noise. Take

$$n = 3, \quad t = 5/4, \quad x = (5/8, 5/8, 0), \quad y = (1, 1/8, 1/8), \quad \theta = t/n.$$

The binomial-noise parameter is $\kappa_y = \sum_j (y_j - \theta)^2/[n\theta(1-\theta)] = 7/10$. Set $x'_i = \theta + \kappa_y(x_i - \theta)$. Then

$$x' = (9/16, 9/16, 1/8), \qquad P_{x'} = (343, 931, 693, 81)/2048, \qquad P_y = (0, 49/64, 7/32, 1/64).$$

At indices $1, 2, 3$, the ratio $P_y/P_{x'}$ is $(1568/931, 64/99, 32/81)$. It is not log-concave, because

$$\frac{(64/99)^2}{(1568/931)(32/81)} = \frac{76}{121} < 1.$$

Mean-interval endpoints. Let T_κ be binomial noise on the count of n independent Bernoulli(θ) bits. Each bit is retained with probability κ; otherwise it is replaced by an independent Bernoulli(θ) bit. Write $\langle f, g\rangle_B = \sum B(k)f(k)g(k)$. For $n = 3$, take both laws to be $P_t = 1 + \mathrm{Bern}(t-1)$, with $1 \le t \le 2$. Set

$$r_t = P_t/B, \qquad \kappa_t = 1 - \frac{\mathrm{Var}(P_t)}{3\theta(1-\theta)}, \qquad D_t = \sqrt{\langle r_t, T_{\kappa_t} r_t\rangle_B} - \langle r_t, r_t\rangle_B.$$

At each endpoint, $D_t = \sqrt{9/4} - 9/4 = -3/4$. At $t = 3/2$,

$$r_t = (0, 4/3, 4/3, 0), \qquad D_t = \sqrt{31/27} - 4/3 > -3/4.$$

Thus the maximum cannot in general be restricted to the two endpoints. This example does not violate the desired inequality $D_t \le 0$.

A.17 A collision probability need not decrease with the trial counts

A more general collision comparison has two proved boundary cases. Let r, s be positive integers, set $v = rs/(r+s)$, and let $h \ge r$, $N \ge s$ be integers. The proposed bound is

$$\Pr\{\mathrm{Bin}(h, v/h) = \mathrm{Bin}(N, v/N)\} \le \binom{r+s}{r} \left(\frac{r}{r+s}\right)^r \left(\frac{s}{r+s}\right)^s. \tag{191}$$

The two binomial variables are independent. The right side is the collision probability at $h = r, N = s$. Bound (191) holds on both edges $h = r$ and $N = s$.

To prove this, put $\gamma = s/(r+s)$. Take independent $R \sim \mathrm{Bin}(h, r/h)$ and $S \sim \mathrm{Bin}(N, s/N)$. Retain each success of R with probability γ, and each success of S with probability $1-\gamma$. The retained counts have the laws on the left of (191). Given R, S, their collision probability is

$$H(R,S) = \binom{R+S}{R} (1-\gamma)^R \gamma^S.$$

This identity follows by summing the two binomial probabilities and using Vandermonde's identity. If $h = r$, then $R = r$ deterministically, and

$$\frac{H(r, S+1)}{H(r,S)} = \gamma \frac{r+S+1}{S+1}.$$

The maximum is attained at $S = s-1, s$, so averaging proves (191). Interchanging the two sides proves the other edge. The interior case is not proved by this argument.

For $r = s = 1$, the endpoint collision comparison asks whether

$$\Pr\{\mathrm{Bin}(h, 1/(2h)) = \mathrm{Bin}(N, 1/(2N))\} \le \frac{1}{2}, \qquad h, N \ge 1,$$

where the two variables are independent. A stronger assertion that the probability decreases in each trial count is false. With equal counts, the probability is

$$L_h = \sum_{k=0}^{h} \binom{h}{k}^2 \left(\frac{1}{2h}\right)^{2k} \left(1 - \frac{1}{2h}\right)^{2h-2k}.$$

Exact evaluation gives

$$L_2 = \frac{59}{128} < \frac{1154188663}{2500000000} = L_5.$$

The limiting value is $e^{-1}\sum_{k\ge 0} 1/[4^k(k!)^2]$, also written $e^{-1}I_0(1)$. Both the finite increase and this limit concern the monotonicity assertion. They do not refute the bound $1/2$.

A.18 Column averaging can increase a rank-two permanent

Let

$$A = \frac{1}{30}\begin{pmatrix} 4 & 15 & 11 \\ 4 & 15 & 11 \\ 22 & 0 & 8 \end{pmatrix}, \qquad \widetilde{A} = \frac{1}{30}\begin{pmatrix} 4 & 13 & 13 \\ 4 & 13 & 13 \\ 22 & 4 & 4 \end{pmatrix}.$$

Both matrices are doubly stochastic and have rank two. The second averages columns two and three, but

$$\operatorname{per}(A) = \frac{137}{450}, \qquad \operatorname{per}(\widetilde{A}) = \frac{689}{2250}, \qquad \operatorname{per}(\widetilde{A}) - \operatorname{per}(A) = \frac{2}{1125} > 0.$$

Hence arbitrary column mixing does not decrease the permanent, even in rank two. The same matrix also satisfies

$$\operatorname{per}(A_{[:,(1,2,2)]}) = \frac{11}{30} > \frac{137}{450}.$$

Thus the scalar likelihood-peak theorem does not give a pointwise maximum at the central multivariate occupancy, even in rank two.

A.19 Minimum Fischer norm cannot supply a proposed deletion factor

For a nonnegative degree-n polynomial in n variables, define

$$\mathcal{C}_F(p) = \inf_{D>0} \frac{N(p(Dx))}{(\det D)^2}, \qquad g_n = \left(\frac{n-1}{n}\right)^{n-1}, \quad g_1 = 1.$$

The proposed comparison

$$\frac{\mathcal{C}_F(p_A)}{N(p_A)} \ge \frac{g_n}{g_d}, \tag{192}$$

where d is the number of positive entries in a zero-containing column, fails even in rank two. Take $n = 50$,

$$u = (0, 1/2, \underbrace{1/96, \dots, 1/96}_{48}), \qquad A_{i1} = u_i, \quad A_{ij} = (1-u_i)/49 \quad (j \ge 2).$$

This doubly stochastic matrix has rank two and exactly one zero entry. Thus $d = 49$. Let

$$P = \mathrm{Bern}(1/2) + \mathrm{Bin}(48, 1/96), \qquad B = \mathrm{Bin}(50, 1/50), \qquad C(z) = \sum_{k=0}^{49} \frac{P_k^2}{B_k} z^k.$$

Writing $Y = x_2 + \cdots + x_{50}$ gives $p_A = \sum_k P_k x_1^k (Y/49)^{50-k}$. The multinomial theorem then gives $N(p_A) = c_{50} C(1)$. Scaling only x_1 by $\sqrt{z}$ has squared determinant z, so

$$\frac{\mathcal{C}_F(p_A)}{N(p_A)} \le \frac{C(z)}{zC(1)}.$$

For a direct rational certificate, put

$$P_k = \frac{\binom{48}{k} 95^{48-k} + \binom{48}{k-1} 95^{49-k}}{2 \cdot 96^{48}}, \qquad B_k = \frac{\binom{50}{k} 49^{50-k}}{50^{50}},$$

with each out-of-range binomial term set to zero before evaluation. Substitution and rational comparison give

$$\frac{C(31/32)}{(31/32)C(1)} < \frac{99973}{100000} < \frac{99979}{100000} < \frac{g_{50}}{g_{49}}.$$

This proves the failure of (192) without an optimizing scale.

The failure persists in all sufficiently large dimensions of the same family. Fix $0 < a < 1$, write $b = 1 - a$, and replace u by $(0, a, b/(n-2), \ldots, b/(n-2))$. The occupancy law tends to $\mathrm{Bern}(a) + \mathrm{Pois}(b)$, while the reference tends to $\mathrm{Pois}(1)$. The limiting collision series is

$$C_\infty(z) = e^{1-2b+b^2 z} \big[b^2 (1 + az)^2 + a^2 z \big].$$

For $1 \le k \le n$, the finite laws satisfy

$$P_{n,k} \le \frac{b^{k-1}(b^2 + ak)}{k!}, \qquad B_{n,k}^{-1} \le e^{k+2} k!.$$

Bound the $k = 0$ term separately. Set the finite collision coefficients to zero for $k > n$. These estimates give summable factorial majorants for the collision series and its derivative on compact z-intervals. Dominated convergence applies. Direct differentiation gives

$$\frac{C'_\infty(1)}{C_\infty(1)} - 1 = \frac{a^4 b^2}{1 - a^2 + a^4} > 0.$$

Thus a fixed scale just below one decreases $C_\infty(z)/z$, while $g_n/g_{n-1} \to 1$. This is the stated dimension-persistent failure. It does not refute the capacity-balanced deletion inequality, the strong Gram inequality, or Gyires's inequality.

A.20 An entropy reduction for scalar reference laws

Let n, m be integers with $1 \le m \le n - 1$, and put $B = \mathrm{Bin}(n, m/n)$. Let P have mean m, and suppose $r = P/B$ is log-concave with interval support. Define

$$D(P\|B) = \sum_k P_k \log r_k, \qquad D_2(P\|B) = \log \sum_k B_k r_k^2.$$

Concavity of the piecewise-linear interpolation of $\log r$ and the mean condition give

$$D(P\|B) \le \log r_m.$$

Consequently $D_2(P\|B) \le 2D(P\|B)$ would imply $\sum_k B_k r_k^2 \le r_m^2$. On a localized interval law, take integers $0 \le a \le m \le b \le n$ and $t > 0$. Set $r_k = t^{k-m}/Z$ for $a \le k \le b$, and set it to zero otherwise. Here $Z = \sum_{k=a}^{b} B_k t^{k-m}$. With the tilt chosen to give mean m, one has $D(P\|B) = -\log Z = \log r_m$. Thus the entropy comparison is equivalent to the norm comparison on those localized laws. This reduction does not prove the comparison for asymmetric integer means outside the established cases.

An induction by division of a common variable has a separate failure in the Lorentzian class. The binary cubic

$$p(X,Y) = Y^3 + 3XY^2 + 3X^2Y$$

is Lorentzian because its coefficients divided by $\binom{3}{k}$ are $(1,1,1,0)$, a log-concave sequence with interval support. But

$$p/Y = Y^2 + 3XY + 3X^2$$

has discriminant -3, so it is not a Lorentzian binary quadratic. Division by a common variable therefore does not preserve Lorentzianity in general. This failure does not apply to division within the real-stable class.

Two proper truncations give concrete norm ratios. For $B = \mathrm{Bin}(n, 1/n)$, set $u_k = t^{k-1}$ for $0 \le k \le L$, and zero otherwise. Choose $t > 0$ so that $\sum_k (k-1) B_k u_k = 0$. Normalizing $B_k u_k$ gives a mean-one law. Its scalar norm ratio is

$$\frac{r_1^2}{\sum_k B_k r_k^2} = \frac{1}{\sum_k B_k u_k^2}.$$

For $n = 3, L = 2$, the mean equation gives $t^2 = 4/3$, and

$$\sum_k B_k u_k^2 = \frac{2}{9} + \frac{4}{9} + \frac{8}{27} = \frac{26}{27}.$$

The ratio is exactly $27/26$. For $n = 4, L = 3$, the tilt is the unique positive root of $8t^3 + 18t^2 - 27 = 0$. The ratio is

$$\frac{256}{81/t^2 + 108 + 54t^2 + 12t^4} \approx 1.003921142546.$$

This value is below $256/255 \approx 1.003921568627451$, but remains above one. The displayed root equation specifies this example exactly. The all-order scalar proof does not depend on this numerical comparison.

A.21 An equivalent product inequality

In each fixed order, universal validity of the strong inequality (187) is equivalent to

$$c_n \operatorname{per}(AB) \le \operatorname{per}(A)\operatorname{per}(B) \tag{193}$$

for all doubly stochastic A, B. Indeed, permanent Cauchy–Schwarz gives

$$\operatorname{per}(AB)^2 \le \operatorname{per}(AA^{\mathsf{T}})\operatorname{per}(B^{\mathsf{T}}B).$$

Apply (187) to A and B^{T}, then take nonnegative square roots to obtain (193). Conversely, use $B = A^{\mathsf{T}}$ in (193). Thus known special families do not automatically give closure under matrix multiplication. That closure would need an additional permanent comparison.

A.22 Tensor forms of the two inequalities

Let $e_1, \dots, e_n$ be the standard orthonormal basis, and define

$$u = n^{-1/2}(1, \dots, 1), \qquad v = \frac{1}{\sqrt{n!}} \sum_{\sigma \in S_n} e_{\sigma(1)} \otimes \cdots \otimes e_{\sigma(n)}.$$

Then $\|v\| = 1$ and $\langle v, u^{\otimes n} \rangle = \sqrt{c_n}$. Expansion of the permutation sums gives

$$\operatorname{per}(A) = \langle v, A^{\otimes n} v \rangle, \qquad \operatorname{per}(AA^{\mathsf{T}}) = \|(A^{\mathsf{T}})^{\otimes n} v\|^2.$$

For $h = (A^{\mathsf{T}})^{\otimes n} v$, double stochasticity implies $\langle h, u^{\otimes n} \rangle = \sqrt{c_n}$. Thus (187) says that the cosine of the angle between h and v is at least $\sqrt{c_n}$. Section A.14 shows why balanced nonnegative coefficient laws alone cannot establish the needed norm bound.

For the original inequality, put $T = A^{\otimes n}$ and $H = (T + T^*)/2$, where T^* is the transpose operator. Direct multiplication gives

$$\langle v, Hv \rangle = \operatorname{per}(A), \qquad \|Hv\|^2 = Q(A).$$

The original inequality is therefore $\langle v, Hv \rangle^2 \ge c_n \|Hv\|^2$. This identity retains both matrix orientations. It is a reformulation, not a proof of the angle bound.

A.23 The first two gap layers for Latin-square weight matrices

Let $n \ge 3$, and let every row and column of B be a permutation of a vector w. Let E be the all-ones matrix, put $s = \sum_i w_i$, and define

$$F(A) = n^n \operatorname{per}(A)^2 - n! \operatorname{per}(AA^{\mathsf{T}}), \qquad V(w) = \sum_{i<j} (w_i - w_j)^2,$$

$$W(w) = \sum_{i<j} (w_i - w_j)^2 (w_i + w_j).$$

The first two possibly nonzero powers in the expansion about tE are

$$\frac{F(tE + B)}{n^n n!(n-2)!} = t^{2n-2} V(w) + t^{2n-3} H_3(w) + \text{terms of degree at most } 2n-4 \text{ in } t, \tag{194}$$

where

$$H_3(w) = \left(\frac{2(n-1)}{n} - \frac{8}{3n(n-2)}\right) sV(w) + \frac{4}{3(n-2)} W(w). \tag{195}$$

Here is a derivation. Set $X = B - (s/n)E$ and $a = t + s/n$. Then X has zero row and column sums. Expanding the permanent by the entries selected from X gives

$$\operatorname{per}(aE + X) = n!a^n + \frac{(n-2)!}{2} a^{n-2} \sum_{ij} X_{ij}^2 + \frac{2(n-3)!}{3} a^{n-3} \sum_{ij} X_{ij}^3 + \cdots .$$

The quadratic and cubic coefficients follow by summing products over distinct rows and columns and then using the zero sums. For the cubic coefficient, inclusion–exclusion for three row labels and three column labels leaves $4 \sum_{ij} X_{ij}^3$ in the ordered sum; division by 3! gives $2 \sum X_{ij}^3/3$.

Also, $(aE+X)(aE+X)^{\mathsf{T}} = na^2E + XX^{\mathsf{T}}$. The term linear in XX^{T} in its permanent vanishes, since the entries of XX^{T} sum to zero. Thus the Gram permanent has no contribution of degree two or three in X. The constant terms cancel in F. After the normalization in (194), the next terms are

$$a^{2n-2} \sum X_{ij}^2 + \frac{4}{3(n-2)} a^{2n-3} \sum X_{ij}^3.$$

Every row of X permutes $w - s/n$, so

$$\sum X_{ij}^2 = V(w), \qquad \sum X_{ij}^3 = W(w) - \frac{2s}{n} V(w).$$

Substitute $a = t + s/n$ and collect the two highest remaining powers of t. This proves (194)–(195).

For nonnegative sorted w, write each coordinate as a sum of consecutive nonnegative gaps. Each squared difference, each sum $w_i + w_j$, and s then has nonnegative coefficients. Both scalar factors in (195) are nonnegative for $n \geq 3$. Therefore the full quadratic and cubic gap layers have nonnegative coefficients, for every Latin-square arrangement. This statement gives no sign control over higher gap degrees.

A.24 A two-product relaxation for balanced stable ternary cubics

This reduction uses the definite determinantal representation theorem for ternary hyperbolic polynomials. A reference is Hanselka, *Definite Determinantal Representations of Ternary Hyperbolic Forms*, Theorem 1, arXiv:1411.1661v2. Let p be a balanced stable ternary cubic with nonnegative coefficients and $p(\mathbf{1}) = 1$. The theorem gives a representation

$$p(x, y, z) = \det(xA + yB + zC), \qquad A, B, C \succeq 0,$$

$$A + B + C = I_3, \qquad \operatorname{tr} A = \operatorname{tr} B = \operatorname{tr} C = 1.$$

Congruence normalizes the matrix at $\mathbf{1}$ to the identity. Differentiating the determinant at the identity gives the traces. The positive orthant lies in the closure of the hyperbolicity cone, so the three coefficient matrices are positive semidefinite.

First assume $I - A$ is positive definite. Congruence by $(I - A)^{-1/2}$, followed by an orthogonal change of basis, gives

$$p = \det(I - A)\det(xR + \operatorname{diag}(\beta_i y + (1 - \beta_i)z)), \quad 0 \le \beta_i \le 1.$$

Let r_1, r_2, r_3 be the eigenvalues of R, and let W contain the squared entries of an orthogonal eigenvector matrix. It is doubly stochastic. In dimension three,

$$R_{ii} = \sum_j W_{ij} r_j, \qquad (\operatorname{adj} R)_{ii} = \sum_j W_{ij} \prod_{k \neq j} r_k.$$

Expansion against the diagonal matrix shows that the full coefficient vector is affine in W. A Birkhoff decomposition therefore expresses p as a convex combination of

$$f_\pi(x, y, z) = \prod_{i=1}^{3} \Big(a_{\pi(i)} x + (1 - a_{\pi(i)})\beta_i y + (1 - a_{\pi(i)})(1 - \beta_i) z \Big), \tag{196}$$

where $a_j = r_j/(1 + r_j)$ are the eigenvalues of A. Thus $0 \le a_j \le 1$ and $\sum_j a_j = 1$. The factor $\det(I - A)$ is absorbed by the three row normalizations.

Every f_π has value one and x-derivative one at $\mathbf{1}$. Its y-derivative there is

$$B_\pi = \sum_i (1 - a_{\pi(i)})\beta_i.$$

The mixture is balanced exactly when its average B_π is one. A hyperplane section of the order-three Birkhoff polytope has vertices at original vertices or at intersections with edges. Each edge joins two permutation matrices whose relative permutation is one cycle. Hence, for fixed a, β, a sufficient convex-cone argument reduces to balanced mixtures

$$(1 - \theta) f_\pi + \theta f_\sigma, \qquad (1 - \theta) B_\pi + \theta B_\sigma = 1,$$

where $\pi^{-1}\sigma$ is a transposition or a three-cycle. Original vertices are included by an endpoint value of θ.

This applies to the norm cone $\{p : [xyz]p \ge \sqrt{c_3}\sqrt{N(p)}\}$, which is convex by the triangle inequality. It also applies to the average-deletion cone $G(p) \ge 0, [xyz]p \ge 0$ from Section A.15, whose defining formula is a Lorentz cone. The individual products in (196) need not be balanced. The relaxed two-product mixtures need not be stable. No assertion that all these mixtures belong to either target cone is proved here.

If $[x^3]p = 0$, then $\det A = 0$, and the eigenvalues can be written $(t, 1 - t, 0)$. If an eigenvalue of A equals one, positivity and $A + B + C = I$ make the other coefficient matrices vanish on its eigenspace. They preserve the orthogonal complement, so the polynomial factors. Limits also cover this endpoint. The determinant reduction does not by itself settle the remaining boundary inequality.

A.25 Two explicit constructions within polynomial deletion

Let $m_1, \dots, m_n$ be nonnegative integers, and put $M = \sum_i m_i \ge n$. Consider M factors, with m_i copies of $x_i + y_1 + \cdots + y_{M-n}$. Extract the coefficient of $y_1 \cdots y_{M-n}$. Choosing the n factors that supply an x-variable and assigning the distinct y-variables to the remaining factors gives

$$[y_1 \cdots y_{M-n}] \prod_{i=1}^{n} (x_i + y_1 + \cdots + y_{M-n})^{m_i} = (M-n)!\, e_n(x_1^{[m_1]}, \ldots, x_n^{[m_n]}).$$

The notation $x_i^{[m_i]}$ means that the argument list contains m_i copies of x_i. This identity places the repeated-variable elementary-symmetric family inside the product-deletion class. It does not prove a norm inequality for that family in arbitrary order.

A second construction has a simpler deletion identity. Let $n \geq 2$ and

$$p = (s + tx_1)^n - (tx_1)^n, \qquad s = x_2 + \cdots + x_n, \qquad t > 0.$$

Every singleton deletion is a positive scalar multiple of a positive linear form raised to power $n-1$. Indeed,

$$[x_1]p = nt\, s^{n-1}, \qquad [x_j]p = n(s - x_j + tx_1)^{n-1} \quad (j \geq 2).$$

Let $b = [x_1 \cdots x_n]p = n!t$. The central coefficient of each deletion is b. A balanced pure power of degree $n-1$ with central coefficient b has squared Fischer norm b^2/c_{n-1}. Hence

$$H([x_j]p) = \frac{b^2}{c_{n-1}}, \qquad \frac{H([x_j]p)}{g_n N(p)} = \frac{b^2}{c_n N(p)} \quad \text{for every } j.$$

These identities remain true after positive variable scaling, with the corresponding central coefficient and norm. Thus this family cannot separate the capacity-balanced deletion comparison from the final coefficient-norm comparison.

A.26 A Hessian criterion and a family of Lorentzian coefficients

Let H be real symmetric, let $e = \mathbf{1}$, and suppose $e^{\mathsf{T}} H e > 0$. The matrix H has at most one positive eigenvalue if and only if

$$\frac{(He)(He)^{\mathsf{T}}}{e^{\mathsf{T}} H e} - H \succeq 0 \quad \text{on } e^{\perp}.$$

To prove the equivalence, decompose a vector into a multiple of e and a vector in $e^{\perp}$. Complete the square in the positive e-coordinate. The remaining quadratic form is the negative of the displayed restricted matrix. Sylvester's inertia law gives the assertion.

This criterion can be applied to a concrete coefficient family. For degree $n \geq 2$ in n variables, take

$$p_\alpha = \frac{\lambda^{\sum_i \binom{\alpha_i}{2}}}{\alpha!}, \qquad 0 < \lambda < 1, \quad |\alpha| = n.$$

For $|\beta| = n-2$, the Hessian of $\partial^\beta p$ is a positive diagonal congruence, with a positive scalar factor, of $E - (1-\lambda)I$, where E is the all-ones matrix. Its diagonal entries before congruence are λ, and its off-diagonal entries are one. It has one positive eigenvalue and $n-1$ negative eigenvalues. The support is the full degree-n simplex, so p is Lorentzian.

One can also set a selected pure coefficient p_{ne_j} to zero. This subtracts a nonnegative diagonal rank-one matrix from the Hessian for $\beta = (n-2)e_j$, and leaves the other Hessians unchanged.

It cannot increase their numbers of positive eigenvalues. The support is the degree simplex with that vertex removed, which is M-convex. The resulting polynomial is therefore Lorentzian as well. This construction supplies admissible examples. It gives no global optimum or norm bound.

A.27 The two Gram permanents need not have a fixed order

A matrix with three exceptional rows need not satisfy $\operatorname{per}(A^{\mathsf{T}}A) \le \operatorname{per}(AA^{\mathsf{T}})$. Take

$$A = \frac{1}{5}\begin{pmatrix} 0 & 3 & 2 & 0 & 0 \\ 2 & 0 & 0 & 3 & 0 \\ 1 & 0 & 1 & 0 & 3 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Every row and column sums to one. Its last two rows equal those of J_5. Direct permanent sums give

$$\operatorname{per}(A) = \frac{42}{625}, \qquad \operatorname{per}(AA^{\mathsf{T}}) = \frac{19926}{390625}, \qquad \operatorname{per}(A^{\mathsf{T}}A) = \frac{508602}{9765625}.$$

Thus

$$\operatorname{per}(A^{\mathsf{T}}A) - \operatorname{per}(AA^{\mathsf{T}}) = \frac{10452}{9765625} > 0.$$

Both strong deficits are positive. They are 624276/244140625 for the row Gram matrix and 15356052/6103515625 for the column Gram matrix. Transposition reverses the displayed strict ordering. A proof for one Gram orientation therefore needs an additional argument for the other orientation.

A.28 Two alternative bounds for the three-shift transfer matrices

Let $C = aP_n^{-1} + bI + cP_n$, where $a, b, c \ge 0$, and define

$$L = \max\left\{a, c, \frac{b + \sqrt{b^2 + 4ac}}{2}\right\}, \quad d = a^2 + b^2 + c^2, \quad e = b(a + c), \quad f = ac.$$

The transfer identity for this family has two scalar blocks f, two blocks with the spectrum of T_1, and one block T_2, where

$$T_1 = \begin{pmatrix} e & d & e & f \\ f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \end{pmatrix}, \qquad T_2 = \begin{pmatrix} d & e & 0 & f & 0 & 0 \\ e & 0 & e & 0 & f & 0 \\ f & 0 & 0 & 0 & 0 & 0 \\ 0 & e & d & 0 & 0 & f \\ 0 & f & 0 & 0 & 0 & 0 \\ 0 & 0 & f & 0 & 0 & 0 \end{pmatrix}.$$

Consequently,

$$\operatorname{per}(CC^{\mathsf{T}}) = 2f^n + 2\operatorname{tr}(T_1^n) + \operatorname{tr}(T_2^n), \qquad \operatorname{per}(C) \ge L^n.$$

The case $L = 0$ is immediate. Otherwise divide the weights by L. Then $0 \le a, c \le 1$ and $b + ac \le 1$. Increasing b to $1 - ac$ increases both transfer matrices entrywise. The following arguments therefore give bounds for all the normalized weights.

An explicit rational comparison vector. Put $K = 51/20$, $k = K - 2 = 11/20$, and $f = ac$. Since $a + c \le 1 + f$,

$$e \le 1 - f^2, \qquad d \le 2 - 2f + 2f^2, \qquad 0 \le f \le 1.$$

For $v = (1, k, K^{-1}, k, kK^{-1}, K^{-2})^{\mathsf{T}}$, the first row of $T_2 v \le Kv$ follows from

$$d + k(e + f) \le K - (2 - k)f(1 - f) \le K.$$

The second row follows from

$$(1 + K^{-1})e + (k/K)f \le \frac{K + 1}{K} + \frac{k^2}{4K(K + 1)} < Kk.$$

The final strict difference is $2881/1448400$. The fourth row is at most

$$k(1 - f^2) + (2 - 2f + 2f^2)/K + f/K^2.$$

Its quadratic coefficient is $239/1020 > 0$. Its maximum on $[0, 1]$ is therefore at an endpoint. The endpoint differences from Kk are $1391/20400$ and $483161/1040400$, both positive. Rows three, five, and six use only $f \le 1$.

For $u = (1, K^{-1}, K^{-2}, K^{-3})^{\mathsf{T}}$, the first row of $T_1 u \le Ku$ follows from

$$1 + 2/K + 1/K^2 + 1/K^3 < K,$$

with difference $1463381/2653020$. The other rows again use $f \le 1$. For any nonnegative matrix T and positive vector v with $Tv \le Kv$,

$$(T^n)_{ii} v_i \le (T^n v)_i \le K^n v_i.$$

After restoring the scale, these bounds give $\operatorname{per}(CC^{\mathsf{T}}) \le 16K^n L^{2n}$. The exact inequalities

$$94^{94} 20^{94} > 16 \cdot 94! \cdot 51^{94}, \qquad 20 \cdot 95^{94} > 51 \cdot 94^{94}$$

show that $16 c_n K^n < 1$ for every $n \ge 94$. Indeed, $(1 + 1/n)^n$ increases with n. This bound proves the tail $n \ge 94$. Its finite range is $3 \le n \le 93$, together with the elementary orders one and two.

A leading-minor certificate with constant $5/2$. With $b = 1 - ac$, form $5I - 2T_1$ and $5I - 2T_2$. Every leading principal minor has positive coefficients in its tensor-product Bernstein basis on $[0, 1]^2$. In increasing minor size, the smallest coefficients are

$$T_1 : \quad 3,\ 37/3,\ 65,\ 325, \qquad T_2 : \quad 1,\ 1,\ 5,\ 25,\ 125,\ 609.$$

These finite comparisons use exact rational arithmetic and require no subdivision. They can be reproduced directly from the displayed matrices. If a minor is $p(a, c) = \sum p_{rs} a^r c^s$ with bidegree (m, h), its Bernstein coefficient at (i, j) is

$$\beta_{ij} = \sum_{r \le i,\ s \le j} p_{rs} \frac{\binom{i}{r}}{\binom{m}{r}} \frac{\binom{j}{s}}{\binom{h}{s}}.$$

The Bernstein basis functions are nonnegative and sum to one. Each minor is therefore strictly positive on the whole square.

The matrices $5I - 2T_j$ have nonpositive off-diagonal entries. Elimination without row swaps has positive pivots, since they are ratios of consecutive leading minors. Every Schur complement retains nonpositive off-diagonal entries. The triangular factors thus have nonnegative inverses. It follows that $(5I - 2T_j)^{-1} \ge 0$. Its positive diagonal gives

$$x = (5I - 2T_j)^{-1}\mathbf{1} > 0, \qquad T_j x = (5x - \mathbf{1})/2 < (5/2)x.$$

The spectral radius is therefore less than $5/2$. The full transfer bound is $16(5/2)^n L^{2n}$. The exact comparisons

$$70^{70}2^{70} > 16 \cdot 70! \cdot 5^{70}, \qquad \frac{5}{2}\left(\frac{6}{7}\right)^6 = \frac{116640}{117649} < 1$$

give the infinite tail from order 70. Its finite range is $3 \le n \le 69$, together with the elementary orders one and two. The sharper constant $1 + \sqrt{2}$ in the main theorem reduces the tail threshold to 48.

The finite ranges. An exact integer coefficient certificate covers all three reflection-reduced sorting chambers for every $3 \le n \le 93$. It checks 273 chambers and 1,674,309 coefficient positions. No coefficient is negative. Reflection exchanges a and c, so these chambers cover all six weight orders. The certificate uses characteristic-polynomial recurrences and Newton identities in sorted-gap coordinates. Run `node independent_finite.mjs 93` to reproduce it. The program writes `independent_finite.json` in the current directory. Compare that output with `results/permutation-alternative-finite.json` in the supporting material. The reproduction guide gives commands that keep the two finite ranges in separate files.

The certificate supplies both finite ranges stated above. Combining it with either tail bound, the elementary orders one and two, and cycle decomposition gives a separate proof for all permutation polynomials $aI + bQ + cQ^2$ with nonnegative weights. This conclusion concerns this three-term family. It does not establish the inequality for general circulants.

B Finite computational observations

This appendix describes stored outputs of finite computations for the permanent and polynomial inequalities. It separates exact arithmetic from numerical optimization. A successful optimization run does not certify a global minimum. A small numerical constraint residual does not certify exact feasibility.

Write $c_n = n!/n^n$, $g_n = ((n-1)/n)^{n-1}$, and

$$\mathcal{R}(A) = \frac{\operatorname{per}(A)^2}{c_n \operatorname{per}(AA^{\mathsf{T}})}.$$

For a balanced degree-n polynomial, write $b = [x_1 \cdots x_n]p$, $N = N(p)$, and let H_j be its capacity-balanced singleton-deletion norm. The three polynomial ratios used below are

$$D_{\mathrm{av}} = \frac{\sum_j H_j}{n g_n N}, \qquad D_{\max} = \frac{\max_j H_j}{g_n N}, \qquad S = \frac{b^2}{c_n N}.$$

The proposed lower bound for each ratio is one. A raw deletion ratio uses $N([x_j]p)$ in place of H_j. All decimal values below are approximate numerical observations, except integers and fractions explicitly identified as exact.

B.1 Exact coefficient tests outside the complete circulant orders

Let $C(w)_{ij} = w_{j-i \bmod n}$, and put $F_n(w) = n^n \operatorname{per}(C(w))^2 - n! \operatorname{per}(C(w)C(w)^{\mathsf{T}})$. For a sorting permutation π, use $w_{\pi(k)} = x_0 + \cdots + x_k$. Each sorting region is called a chamber. The following selected chambers were checked by exact integer polynomial arithmetic.

Order	Checked affine chamber representatives	All affine representatives	Positive coefficients in each checked chamber	Zero coefficients	Least positive coefficient
8	16	1,260	245,149	8	16,736,896
9	5	6,720	1,562,266	9	387,057,609

Every checked coefficient was nonnegative. These are exact statements about the listed chambers. Neither full order was checked.

At order twelve, eight restrictions used the baseline plus two nested gap directions. Each restricted polynomial had 322 nonzero terms and no negative coefficient. The smallest coefficient over these restrictions was 8,915,621,446,656. A three-variable restriction does not certify its full chamber.

A noncyclic Latin-square array of order six was also checked in 50 selected sorting orders. Its rows, using indices zero through five, were

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 1 & 2 & 5 & 0 \\ 1 & 0 & 5 & 4 & 3 & 2 \\ 2 & 3 & 0 & 5 & 1 & 4 \\ 4 & 5 & 3 & 0 & 2 & 1 \\ 5 & 2 & 4 & 1 & 0 & 3 \end{pmatrix}.$$

No checked coefficient was negative. The smallest positive coefficient was 42,624. This finite test does not establish an all-Latin-square theorem.

Additional exact tests used arbitrary lists of permutation matrices. Thirty lists of four permutations at order five, ten lists of six at order six, and thirty lists of four at order eight passed their recorded sorting checks. At order three, all sorting orders of the list of all six permutation matrices passed. The sampled lists do not prove the general coefficient assertion. The exact order-four negative coefficient in the main text refutes that assertion.

B.2 Wider finite checks for the three-shift family

For $C = aP_n^{-1} + bI + cP_n$, exact transfer-polynomial expansion checked all six weight orders for every $5 \le n \le 100$. Direct permanent expansion checked the six chambers at orders three

and four. It also checked both transfer identities in orders one through seven.

A distinct implementation used characteristic-polynomial recurrences and Newton identities directly in gap coordinates. It checked 201 reflection-reduced chambers in orders three through sixty-nine, with 693,249 coefficient positions. No coefficient was negative. At orders four through sixty-nine, every checked coefficient was positive. The overlapping minima agreed with the transfer expansion.

A separate integer recurrence checked all seven nonzero binary triples in orders three through one hundred. There were 686 values. Every deficit at orders four through one hundred was positive. These checks establish finite coefficient and strictness statements for this three-shift family. They do not cover arbitrary circulant weights. The main proof uses only the finite range required by its analytic bound. Appendix A records the separate exact certificate through order ninety-three that supplies both alternative transfer arguments.

B.3 Matrix ratio searches

Constrained local minimization tested $\mathcal{R}(A)$ from 30 initial doubly stochastic matrices in each of orders five, six, and seven. All 90 runs reported convergence. Each order had a smallest recorded ratio of one, within numerical precision. Permanents and gradients used a positive-term subset dynamic program. Finite differences checked the gradient, and direct permutation enumeration checked one permanent.

A separate diagnostic tested 1,000 admissible centered rank-two pairs at each of eight orders:

$$3, 4, 5, 6, 8, 10, 15, 20.$$

It used elementary-symmetric recurrences in extended floating precision. No violation of the strong rank-two target or the proposed duplication-to-Gram bound was reported. The input data, output values, and checking program are not available. These 8,000 reported cases cannot be reproduced from the retained data.

Sparse regular matrices were sampled as sums of disjoint permutation matrices. A positive-term subset dynamic program evaluated the permanents in extended floating precision. The stored aggregate observations are

Order	Degree	Samples	Approximate range of $\mathcal{R}(A)$
16	3	30	16.0–24.3
20	3	20	37.0–54.5
20	4	20	23.4–26.7
20	7	20	5.38–5.43

.

These tests found no counterexample. They were not exhaustive. Individual sampled matrices and detailed run outputs are unavailable. The ratios are numerical observations, not exact bounds. The data directory includes the numerical sampling program `sparse_regular_sampler.cpp`.

B.4 Circulant ratio and Fourier-phase searches

Full-simplex searches used analytic gradients and sequential least-squares programming, with $w \geq 0$, $\sum w_i = 1$, and at most 300 iterations. Starts included Dirichlet weights, sparse supports, uniform weights on random supports, and dense perturbations.

Order	Starts	Convergence reports	Smallest $\mathcal{R}$
8	40	40	$1 - 3.6 \times 10^{-15}$
10	40	40	$1 + 7.1 \times 10^{-15}$
12	40	40	$1 + 3.6 \times 10^{-15}$
14	40	39	$1 - 7.1 \times 10^{-15}$
16	40	38	1

The values slightly below one occur near uniform weights and are at floating-point precision. They are not exact counterexamples.

Fixed-Gram searches kept the Fourier magnitudes and the real Nyquist eigenvalue fixed. They varied conjugate Fourier phases subject to nonnegative inverse-transform weights.

Order	Starts	Convergence reports	Smallest feasible ratio	Smallest ratio among converged runs
8	30	26	1.000000	1.000000
10	30	27	1.120158	1.120158
12	30	25	1.312790	1.312790
14	30	22	1.096436	1.096436
16	30	16	1.300263	2.533075

Feasibility allowed minimum weight -10^{-7}. Tiny negative weights were clipped for final ratio evaluation. Gram errors before clipping were below 1.1×10^{-14}. The distinction between feasible outputs and converged outputs matters, especially at order sixteen. Initial and final weights and numerical residuals are retained for these runs.

B.5 Configuration-model expectations

For the uniform bipartite regular configuration multigraph of degree d, the exact coefficient formulas in the main text evaluated

$$\frac{c_n \, \mathbb{E}[\mathrm{per}(GG^{\mathsf{T}})]}{\mathbb{E}[\mathrm{per}(G)^2]}.$$

Integer arithmetic checked all $1 \le n \le 100$ for $d = 2, 3, 4$, and all $1 \le n \le 60$ for $d = 5, 10, 20$. All 480 ratios were at most one. Equality occurred at order one. Selected numerical displays of these exact ratios are

d	n	Ratio
2	100	1.5610×10^{-7}
3	100	4.8180×10^{-10}
4	100	2.3821×10^{-9}
5	60	4.1872×10^{-5}
10	60	0.0040689
20	60	0.0576722

Exhaustive stub-bijection enumeration independently checked the two expectation formulas at $(d, n) = (2, 2), (3, 2), (2, 3)$. The respective counts were 24, 720, and 720. All sums agreed exactly. An expectation comparison does not give a pointwise inequality for every graph.

A separate stationary-point calculation used a Bethe approximation for the colored matching model at degrees two through twenty. It compared an estimated Gram logarithmic growth rate with $1 + 2[(d-1)\log(d-1) - (d-2)\log d]$. It included a zero-message branch and six positive initial values for root finding. The largest recorded excess was approximately -0.048323685. This is a numerical approximation. It does not establish the true limiting growth rate or a matrix inequality.

B.6 Stable-polynomial and matroid deletion tests

One suite had 52 examples at each degree three through six. At each degree, 40 were products of elementary-symmetric polynomials after a common doubly stochastic substitution. Twelve were determinant polynomials from balanced positive-semidefinite tuples. The initial coefficients and balance were exact; deletion balancing was numerical.

Degree	Examples	Smallest D_{av}	Smallest raw average ratio
3	52	1.0000000	1.0000000
4	52	1.0373282	1.0375230
5	52	1.0242451	1.0242792
6	52	1.0218714	1.0218926

The largest balancing-gradient residual was below 4.5×10^{-8}. All 208 recorded final ratios S were at least one.

A second suite had 90 examples per degree. Sixty used elementary-symmetric factors on different positive linear forms. Thirty used generic positive-semidefinite determinant pencils. Initial and deletion balancing used convex log-sum-exp minimization.

Degree	Examples	Smallest D_{av}	Smallest raw average ratio
3	90	1.0198150	1.0198288
4	90	1.0180926	1.0181045
5	90	1.0204552	1.0204734
6	90	1.0251072	1.0251346

The largest residual was below 2.4×10^{-10}. The ratio of the scaled deletion sum to the raw sum ranged from approximately 0.95592 to 0.999997. All 360 final ratios S passed. Determinant coefficients were obtained by integer finite differences with Bareiss elimination.

A third suite substituted positive linear forms into matroid basis polynomials, then balanced the result. Its 750 examples had the following minima.

Source matroid	Degree	Examples	Smallest D_{av}	Smallest S
Fano	3	150	1.0091356	1.0181519
Dual Fano	4	150	1.0213981	1.0648370

Source matroid	Degree	Examples	Smallest D_{av}	Smallest S
Fano plus one coloop	4	150	1.0167878	1.0513456
Fano plus two coloops	5	150	1.0190338	1.0782817
Fano plus three coloops	6	150	1.0149026	1.0762892

These are Lorentzian inputs. The Fano source is not stable, although a substitution can produce a stable polynomial. Every raw average passed. The largest residual was below 4.1×10^{-9}. Together, the three suites contain 1,318 saved examples. They do not prove a stable or Lorentzian deletion theorem.

Nine coefficient patterns obtained by grouping the seven Fano variables into three groups of at least two variables were also reported to pass numerical average deletion. No separate output record for those nine patterns is available. An additional 800 matrix-product tests, 200 at each order four through seven, reported no violation. Their individual inputs were not retained.

B.7 Circulant deletion tests

A coefficient-array implementation tested eight optimization starts at each order four through seven. All 32 runs reported convergence. The respective minima of D_{av} were approximately 1, $1 + 8.0 \times 10^{-11}$, $1 + 1.7 \times 10^{-12}$, and $1 + 5.3 \times 10^{-12}$. Residuals were below 8.5×10^{-13}.

A permanent-gradient implementation agreed with the coefficient-array method within 10^{-14} on 30 random cases in orders three through seven. It then tested 30, 30, 20, 9, and 6 starts at orders eight, ten, twelve, fourteen, and sixteen. All 95 runs reported convergence. No ratio below one beyond numerical precision was found. Capacity-gradient residuals were below 2.8×10^{-12}. This numerical evidence is separate from the proved two-shift deletion theorem.

B.8 Nested singleton deletion and concentrated elementary-symmetric inputs

Nested tests started from a product of $d + r$ positive linear forms in $d + r$ variables. They extracted exponent one from r variables, then balanced the residual degree-d polynomial. Coefficient extraction used dynamic programming. Differentiation through capacity minimizers used covariance Hessians. Three small integer kernels agreed with naive coefficient expansion exactly. Directional gradient checks had errors below 1.5×10^{-9}.

Average-deletion optimization saved 43 runs: four at $(d, r) = (3, 3)$, seven at $(3, 8)$, sixteen at $(4, 6)$ across two parameterizations, eight at $(6, 4)$, and eight at $(8, 2)$. Forty-two reported convergence. No saved average or unrestricted maximum ratio was below one. Best-point residuals were below 2.0×10^{-12}.

Prescribed-coordinate tests forced $r + 1$ factors to omit one variable, so its residual pure coefficient vanished. Twelve runs at $(3, 5)$ had smallest prescribed ratio approximately 1.11622. One run each at $(4, 0), (5, 3), (6, 2)$ gave approximately 0.95461005, 0.85057133, and 0.77161423. The average and unrestricted maximum ratios stayed above one. The exact order-four counterexample in the main text establishes failure of the prescribed rule. The degree-five and degree-six values remain numerical observations.

A compressed two-class calculation tested repeated-variable elementary-symmetric polynomials. Balance reduced to one noncentral-hypergeometric mean equation. Norms and deletions reduced to univariate coefficient products. Infinite multiplicity used a Poisson limit; multiplicity one used a balancing limit.

Degree	Cases	Smallest D_{av}
3	16	1.0033444425
4	24	1.0025067642
8	40	1.0012531916
16	48	1.0006267742
32	48	1.0003134584
64	48	1.0001567502
128	48	1.0000783808
256	48	1.0000391918

All 320 average, maximum, and final ratios passed numerically. The largest mean residual was approximately 5.12×10^{-12}. Four small finite examples agreed with full multivariate expansion within 6×10^{-14}. These tests do not prove an all-degree result.

B.9 Direct Lorentzian coefficient optimization

Cubic searches normalized the central coefficient and constrained balance, nonnegative coefficients, and derivative-Hessian signatures. Of 160 outputs, 134 passed the stated numerical feasibility screen: 63 of 80 deletion-objective outputs and 71 of 80 norm-objective outputs. The smallest recorded cubic deletion deficit $G = 3\sum_j H_j - 4N$ among these outputs was approximately -3.76×10^{-8}. A numerical screen near a singular Hessian does not certify Lorentzianity or an exact negative deficit.

A boundary cubic search retained 62 outputs, of which 53 passed its screen. Its smallest G was approximately $2/3$. A separate mixed-coefficient search retained 100 outputs, of which 91 passed. Its largest mixed sum was approximately $10/3$. The exact polynomial in Appendix A realizes these latter two values. The search does not prove the proposed universal boundary bound $G \geq 2b^2/3$.

At degree four, a 35-output boundary search had eight outputs that passed its numerical signature screen, but none reported optimizer convergence. Its smallest screened final ratio was approximately 1.003921125. A distinct eight-start direct constrained search had no feasible output. Those rejected outputs are not counterexamples.

A Schur-matrix barrier method then kept accepted iterates in the strict numerical Hessian region. It enforced balance through an affine nullspace. Four degree-four starts reached a smallest average ratio of approximately 1.003923425. Two degree-five starts reached approximately 1.00032812 and 1.00033093. The coefficient counts were 35 and 126; the Hessian counts were 10 and 35. Best-point balance residuals were below 3×10^{-13}. The degree-four deletion-capacity residual reached 2.05×10^{-10} in one saved best point. These points have numerical feasibility checks only.

An additional cubic optimization was reported to have only nine feasible outputs among 120. Those nine gave no violation. Its underlying output file is not available, so its detailed constraints and values cannot be checked from a saved run.

Tests of fixed low-degree sum-of-squares forms used nonnegative combinations of Alexandrov–Fenchel minors, derivative-Hessian determinants, balance identities, and related reverse-Cauchy

terms. Stable-only variants included Laguerre inequalities. The tested forms were reported infeasible. No exact dual certificate is supplied. This does not prove that other certificates are impossible. A numerical probe of the two-product determinantal relaxation also reported no failure, but no completed run record with counts and input values is available.

B.10 M-convex support enumeration and attenuation

For an M-convex support T, let

$$F_T(z) = \sum_{\alpha \in T} \frac{c_n}{\alpha!} e^{(\alpha - \mathbf{1}) \cdot z}, \qquad \nabla F_T(z_*) = 0.$$

The target is $F_T(2z_*) \leq 1$. A boundary support uses the corresponding balancing limit.

At degree three, exact subset enumeration and exchange checks found 57 labeled supports containing the center, with 17 coordinate-permutation orbits. All numerical target ratios passed. The largest balancing residual was approximately 1.67×10^{-12}.

At degree four, the search covered all full-dimensional integral-polymatroid supports with the center in the interior. Of 5,184 candidate rank assignments, 615 were monotone and submodular. They gave 63 support orbits. Every support passed an exact exchange check. Numerical balancing gave no target violation. The largest proper-support value was approximately 0.9960941728. The largest residual was approximately 9.63×10^{-12}. All 45,456 strict inclusion pairs among the 615 labeled supports were compared. No numerical inclusion-monotonicity violation was found.

A nonexhaustive degree-five beam search had width nine and a requested depth of 45. It completed depth levels one through thirty-three, visited 636 supports, and checked 1,537 rank-step inclusions. Rank and submodularity checks were exact. The largest proper-support value was approximately 0.9996800006. No target or inclusion failure was found. The largest state residual was approximately 1.51×10^{-8}.

Fifty attenuation paths from asymmetric supports were sampled at seven interior parameters each. The largest interior value was approximately 0.6914288203. There were 252 samples above the endpoint chord. The largest interior balancing residual was approximately 1.33×10^{-12}. The exact nonconvexity example in the main text establishes failure of continuous convexity independently of these numerical values.

The duplication comparison was also evaluated on the same 17, 63, and 636 supports. The complement-pair coefficient

$$d_T = \sum_{\substack{\alpha \in T \\ 2\mathbf{1} - \alpha \in T}} \frac{1}{\alpha!(2\mathbf{1} - \alpha)!}$$

was exact. The denominator used numerical balancing. Every observed ratio passed, with minimum one at the full support. This does not prove a general duplication bound.

B.11 Rank-two scalar and deletion tests

A truncated-binomial scalar probe covered orders two through eighteen and every half-integer mean strictly between zero and the order. This gives 323 order-mean pairs. At each pair it enumerated admissible support intervals and solved each tilt equation numerically. It found no ratio below $1 - 10^{-8}$. The smallest recorded ratio was 0.9999999999999953, at order eighteen

and mean 5/2. The saved report contains the minimum and failures, not every interval-pair value. The associated broad ultra-log-concave assertion is not established by this probe; the exact obstruction in the main text must retain its stated scope.

For the shifted-binomial profile comparison, let $B = \mathrm{Bin}(n, t/n)$, with $0 < t < n$. Let P, Q be shifted-binomial profile laws on $\{0, \dots, n\}$, both with mean t. Write $r = P/B$ and $s = Q/B$ for their likelihood ratios. These laws satisfy $0 \le \mathrm{Var}(Q) \le t(1 - t/n)$. The tested ratio was

$$\frac{\langle r, s\rangle_B^2}{\langle r, T_{\kappa(s)} r\rangle_B}, \qquad \kappa(s) = 1 - \frac{\mathrm{Var}(Q)}{t(1 - t/n)}.$$

Here T_κ is the binomial count-noise operator. It keeps each Bernoulli coordinate with probability κ and independently resamples it otherwise.

A numerical search of these profiles covered orders sixteen through twenty-two. It tested 175,574 ordered profile pairs and 5,723,130 initial trial means, followed by local one-dimensional refinements. It found no negative contraction deficit. The arithmetic was floating point. The trial schedule included points close to each endpoint and points inside each unit mean interval. Identical profiles and full-binomial profiles were omitted for separate mathematical reasons. This search does not extend the exact rank-two theorem beyond order fifteen.

Ten saved profile examples also violate a proposed endpoint-dominance shortcut. The exact interior-mean example in Appendix A establishes that failure. A duplication-bridge diagnostic had a smallest recorded deficit of approximately -5.55×10^{-16}, at a full-binomial equality case. This value does not establish a strict negative deficit.

A rank-two deletion search saved an arbitrary-column failure at order three, with ratio approximately 0.96423013. The exact rational matrix in the main text proves that failure. Zero-containing-column searches saved extremizers at orders three through eight. Their smallest selected ratios were approximately

$$1.12017221,\ 1.06856934,\ 1.05402263,\ 1.03561156,\ 1.02709402,\ 1.02357208.$$

A separate profile search took the maximum over the eligible zero-containing columns. Its six saved values, in the same orders, were approximately

$$1.35637345,\ 1.12805402,\ 1.10598346,\ 1.06057078,\ 1.04883878,\ 1.04642169.$$

The saved extremizers and residuals permit inspection of those outputs. They are not a record of every sampled matrix and do not prove the zero-containing-column deletion assertion.

B.12 Stochastic complements and minimum-norm deletion

For a pivot $a_{ij} < 1$, define the stochastic complement

$$B^{(ij)}_{k\ell} = a_{k\ell} + \frac{a_{kj} a_{i\ell}}{1 - a_{ij}} \quad (k \ne i,\ \ell \ne j),$$

and let $R(A) = \mathrm{per}(A)^2 / \mathrm{per}(AA^{\mathsf{T}})$. The diagnostic pivot ratio is $q_{ij} = g_n R(B^{(ij)})/R(A)$.

An initial matrix search retained extrema in orders three through seven. For each completed order, its program sampled 90 integer sums of permutation matrices. It tested duplication on the first 25 and tested all eligible pivots. The smallest duplication ratios ranged from

1.02143277 to 1.16495567. The largest observed minimum-pivot ratios were approximately 0.99540045, 0.89612551, 0.84375269, 0.87753059, and 0.89692640. Only the extrema were saved. The programmed order-eight run has no completed output.

A further pivot search retained order-three through order-seven summaries, with a schedule of 160 matrices per order. It examined minima and averages weighted uniformly, by entries, by complementary entries, by matching marginals, and by their complements. Some averages exceeded one. The largest uniform average among the retained summaries was approximately 1.10847118. The largest complement-of-matching average was approximately 1.39060221. The saved summaries contain their extremizing matrices. They do not establish the existence of a successful pivot for every matrix.

Direct constrained optimization maximized $n^{-1}\sum_{ij} a_{ij}q_{ij}$. It saved 24 converged runs in each of orders three, four, and five, and 14 at order six. The largest scores were approximately one in the first three orders and 1.014299626 at order six. Constraint residuals were below 3.6×10^{-13}. The exact stochastic-complement counterexamples in the main text establish the relevant failures without numerical optimization.

A separate search minimized the average of determinant-normalized minimum Fischer norms of deletions. It retained one extremizer in each order three through seven, from a schedule of 180 samples per order. The five average ratios were approximately 1.02219762, 1.02482586, 1.04882773, 1.03986009, and 1.03887917. Maximum balancing residuals were below 7.1×10^{-10}. The exact order-700 example in the main text refutes this averaged minimum-norm comparison. These smaller numerical passes do not limit that counterexample.

B.13 Finite arithmetic checks of analytic formulas

Exact rational diagnostics checked 360 three-exceptional-row matrices in orders five through ten, and 51 at orders $31, 32, 33, 40, 64$. The latter used a rank-three permanent expansion, checked against subset dynamic programming in orders five through seven. All tested margins passed. The all-order conclusions depend on the analytic proofs, not these samples.

Additional exact checks covered 405 integer-mean Poisson-binomial laws in orders two through ten, 28 two-column matrix identities, 184 four-block matrices in orders two through eight, and six row-pair affine identities. All passed. Separate local polynomial calculations checked the first and second variation formulas in degrees two through six for rank-one and full-rank projection directions. Those variation checks used numerical balancing and do not prove a neighborhood theorem.

For the finite rank-two certificate, a rational power-basis implementation agreed with an integer homogeneous implementation on all 454 profile intervals in orders two through six, comprising 20,452 coefficients. Direct permanent calculations tested 378 profile pairs in orders two through seven, with 1,390 identity comparisons, 21 singleton cases, and 251 uniform-profile instances. A further check supplied 209 direct comparisons. These tests support formula implementation. The exact whole-interval coefficient certificate and the analytic extremal reduction establish the finite theorem.

B.14 Data limits

The accompanying scientific data contain 69 computational JSON records and one JSON extraction from a numerical profile log. The extraction preserves every recorded numerical point and each completed-order count, including the printed decimal precision. These records retain input parameters or extremizing inputs, objective values, and residuals when those

fields were available. Summary-only records cannot reconstruct every sampled point. The unavailable items identified above are the individual centered rank-two samples, the 800 additional matrix-product samples, the nine Fano-identification outputs, the 120-output cubic feasibility run, the sparse high-order matrix outputs, and completed output records for the two-product relaxation and fixed sum-of-squares probes.

The absence of those records does not alter an exact proof given in the main text. It limits what can be reproduced about the associated finite observation. No finite numerical pass, failed optimization, or unavailable diagnostic is used to assert the unrestricted Gyires inequality.

References

[1] Darcy Best and Ian M. Wanless. What did Ryser conjecture?, 2018. URL https://arxiv.org/abs/1801.02893. https://doi.org/10.48550/arXiv.1801.02893.

[2] Julius Borcea, Petter Brändén, and Thomas M. Liggett. Negative dependence and the geometry of polynomials. *Journal of the American Mathematical Society*, 22(2):521–567, 2009. doi: 10.1090/S0894-0347-08-00618-8. URL https://arxiv.org/abs/0707.2340.

[3] Petter Brändén and June Huh. Lorentzian polynomials. *Annals of Mathematics*, 192(3):821–891, 2020. doi: 10.4007/annals.2020.192.3.4. URL https://annals.math.princeton.edu/2020/192-3/p04.

[4] Derek K. Chang. On two permanental conjectures. *Linear and Multilinear Algebra*, 26(3):207–213, 1990. doi: 10.1080/03081089008817977. https://doi.org/10.1080/03081089008817977.

[5] Gi-Sang Cheon and Ian M. Wanless. An update on Minc's survey of open problems involving permanents. *Linear Algebra and its Applications*, 403:314–342, 2005. doi: 10.1016/j.laa.2005.02.030. https://doi.org/10.1016/j.laa.2005.02.030.

[6] Leonardo de Moura and Sebastian Ullrich. The Lean 4 theorem prover and programming language. In *Automated Deduction – CADE 28*, volume 12699 of *Lecture Notes in Computer Science*, pages 625–635. Springer, 2021. doi: 10.1007/978-3-030-79876-5_37. https://doi.org/10.1007/978-3-030-79876-5_37.

[7] G. P. Egorychev. The solution of van der Waerden's problem for permanents. *Advances in Mathematics*, 42(3):299–305, 1981. doi: 10.1016/0001-8708(81)90044-x. https://doi.org/10.1016/0001-8708(81)90044-x.

[8] D. I. Falikman. Proof of the van der Waerden conjecture regarding the permanent of a doubly stochastic matrix. *Mathematical Notes of the Academy of Sciences of the USSR*, 29(6):475–479, 1981. doi: 10.1007/bf01163285. English translation of Matematicheskie Zametki 29 (1981), 931–938. https://doi.org/10.1007/bf01163285.

[9] Dev D. Goyal. Gyires's permanent inequality for circulant matrices of order five, 2026. URL https://github.com/dvgyl/gyires-circulant-5/tree/72cf49a51be42cf15592f91d991f0b0c76f58448. Immutable source snapshot; commit 72cf49a51be42cf15592f91d991f0b0c76f58448. https://github.com/dvgyl/gyires-circulant-5/tree/72cf49a51be42cf15592f91d991f0b0c76f58448.

[10] B. Gyires. On permanent inequalities. In A. Hajnal and V. T. Sós, editors, *Combinatorics (Proceedings of the Fifth Hungarian Colloquium, Keszthely, 1976)*, volume 18 of *Colloquia Mathematica Societatis János Bolyai*, pages 471–484. North-Holland, 1978.

[11] B. Gyires. The common source of several inequalities concerning doubly stochastic matrices. *Publicationes Mathematicae Debrecen*, 27(3-4):291–304, 1980. doi: 10.5486/pmd.1980.27.3-4.15. https://doi.org/10.5486/pmd.1980.27.3-4.15.

[12] Wassily Hoeffding. On the distribution of the number of successes in independent trials. *The Annals of Mathematical Statistics*, 27(3):713–721, 1956. doi: 10.1214/aoms/1177728178. https://doi.org/10.1214/aoms/1177728178.

[13] Vladimir Kostov and Boris Shapiro. On the Schur–Szegő composition of polynomials. *Comptes Rendus Mathématique*, 343(2):81–86, 2006. doi: 10.1016/j.crma.2006.06.007. https://www.numdam.org/item/10.1016/j.crma.2006.06.007.pdf.

[14] H. Minc. On permanents of circulants. *Pacific Journal of Mathematics*, 42(2):477–484, 1972. doi: 10.2140/pjm.1972.42.477. https://doi.org/10.2140/pjm.1972.42.477.

[15] Chris Ozols. Gyires' permanent inequality for doubly stochastic matrices of order four, 2026. URL https://github.com/cozuya/gyires-permanent-4x4/tree/244f48c. Lean formalization and documentation; source commit 244f48c. https://github.com/cozuya/gyires-permanent-4x4/tree/244f48c.

[16] The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs*, pages 367–381, 2020. doi: 10.1145/3372885.3373824. Formal development uses Mathlib commit 56a005855ea17e0b0958f73ec801ae140b0a1679. https://doi.org/10.1145/3372885.3373824.