hexagon:2611.00201v1

The Kopotun Subpermanent Transform Inequality

Yair Lavi

11 October 2026

Abstract

We prove the subpermanent transform inequality conjectured by Kopotun in 1996: for every doubly stochastic matrix A of order $n \geq 2$ and every $2 \leq k \leq n$,

$$\sigma_k(A) \geq \sigma_k\left(\frac{nJ_n + A}{n+1}\right).$$

Here σ_k is the sum of all subpermanents of order k, and J_n has every entry $1/n$. Equality holds exactly at $A = J_n$.

Keywords. Permanent; subpermanent sum; doubly stochastic matrix; matrix inequality.
2020 Mathematics Subject Classification. 15A15, 05A20, 15B51.

1 Introduction and main results

For a positive integer r, write $[r] = \{1, \ldots, r\}$ and let S_r be the set of permutations of $[r]$. The permanent of an $r \times r$ matrix $D = (d_{ij})$ is

$$\operatorname{per} D = \sum_{\pi \in S_r} \prod_{i=1}^{r} d_{i,\pi(i)}. \tag{1.1}$$

If D is an $n \times n$ matrix and $S, T \subseteq [n]$ have the same cardinality, let $D[S,T]$ be the submatrix with those rows and columns, listed in increasing order. For $1 \leq k \leq n$, define

$$\sigma_k(D) = \sum_{\substack{S,T \subseteq [n] \\ |S| = |T| = k}} \operatorname{per} D[S,T], \qquad \sigma_0(D) = 1. \tag{1.2}$$

Thus $\sigma_n(D) = \operatorname{per} D$.

A real matrix is **doubly stochastic** if its entries are nonnegative and every row and column sums to one. Denote the set of these $n \times n$ matrices by Ω_n. Throughout, $n \geq 2$, $\mathbf{1}$ is the all-ones column vector, and

$$J_n = \frac{1}{n}\mathbf{1}\mathbf{1}^{\mathsf{T}}, \qquad \tau_n(A) = \frac{nJ_n + A}{n+1} = J_n + \frac{A - J_n}{n+1}. \tag{1.3}$$

The map τ_n takes Ω_n into itself, since it is a convex combination of J_n and A.

Kopotun proposed the inequality $\sigma_k(\tau_n(A)) \leq \sigma_k(A)$ for every $A \in \Omega_n$ and $2 \leq k \leq n$ in 1996 [1, Conjecture 2, p. 167]. Since $\sigma_n = \operatorname{per}$, the case $k = n$ is Wang's permanent conjecture [6], as Kopotun notes [1, Remark 1, p. 167]:

$$\operatorname{per} A \geq \operatorname{per}\left(\frac{nJ_n + A}{n+1}\right).$$

We recently proved this permanent case [5]. Kopotun also recorded the subpermanent inequality for every $2 \le k \le 4$ and every order $n \ge k$ [1, Remark 1], with the endpoint cases $(n,k) = (3,3)$ and $(4,4)$ due to Wang [6] and Chang [7], respectively. Kopotun observed [1, Remark 1] that the transform conjecture follows from the stronger Holens–Đoković conjecture; Wanless later disproved the latter [8]. His paper studies stronger convexity and monotonicity questions as well; the assertion proved here is the specified transform comparison.

Theorem 1.1. *For every integer $n \ge 2$, every $A \in \Omega_n$, and every integer k with $2 \le k \le n$,*

$$\sigma_k(A) \ge \sigma_k(\tau_n(A)). \tag{1.4}$$

Equality holds if and only if $A = J_n$.

We prove the theorem by comparing a quantitative lower bound for $\sigma_k(A)$ with an upper bound for $\sigma_k(\tau_n(A))$. To state these bounds, define the Frobenius norm by $\|X\|_F^2 = \sum_{i,j} x_{ij}^2$ for $X = (x_{ij})$, and put

$$\begin{aligned} S_{n,k} &= k!\binom{n}{k}^2 n^{-k} = \sigma_k(J_n), \\ \alpha_{n,k} &= \frac{k(k-1)}{n(n-1)}, \qquad \Delta_{n,k}(A) = \alpha_{n,k}\|A - J_n\|_F^2, \\ C_n &= \frac{7}{12n^2}, \\ b_n &= \frac{1}{n-1}\left[\left(\frac{n}{n+1}\right)^n \sum_{r=0}^{n} \frac{1}{r!} - 1\right]. \end{aligned} \tag{1.5}$$

The expression for b_n comes from evaluating the permanent at the transform of the identity matrix; see (6.5)–(6.6).

Theorem 1.2. *For every $n \ge 2$, $A \in \Omega_n$, and $2 \le k \le n$,*

$$\frac{\sigma_k(A)}{S_{n,k}} \ge \exp\bigl(C_n \Delta_{n,k}(A)\bigr), \qquad \frac{\sigma_k(\tau_n(A))}{S_{n,k}} \le 1 + b_n \Delta_{n,k}(A). \tag{1.6}$$

Moreover, $0 < b_n < C_n$ for every $n \ge 2$.

Friedland proved the lower bound $\sigma_k(A) \ge S_{n,k}$ [9]. The first estimate of Theorem 1.2 is a quantitative strengthening of that bound.

In particular, if $D_{n,k}(A) = \sigma_k(A) - \sigma_k(\tau_n(A))$ and $\Delta = \Delta_{n,k}(A)$, then

$$\begin{aligned} D_{n,k}(A) &\ge S_{n,k}\bigl(e^{C_n\Delta} - 1 - b_n\Delta\bigr) \\ &\ge S_{n,k}(C_n - b_n)\alpha_{n,k}\|A - J_n\|_F^2. \end{aligned} \tag{1.7}$$

All constants are explicit; they are not asserted to be optimal.

The main step in the lower bound is to represent $\sigma_k(A)$ as a coefficient of a polynomial of degree n. Gurvits's coefficient bound [2] then applies. Its one-variable proof also gives a gain when the roots are unequal. The difficulty is that these roots change when the other variables change. For $k < n$, the polynomial contains $n - k$ uniform linear factors; the arithmetic-geometric mean inequality for those factors controls this movement and preserves enough of the gain. At $k = n$, a direct one-column estimate supplies the lower bound.

Section 2 gives the coefficient inequality and its scalar refinement. Sections 3 and 4 prove the lower estimate in (1.6). Section 5 states the centered subpermanent estimate used for the upper bound and verifies its hypotheses. Section 6 derives that bound, separates the constants, and completes the equality argument. The proof is analytic and does not depend on finite computation.

2 Coefficients and capacity

A polynomial $p \in \mathbb{R}[z_1, \dots, z_r]$ is **real stable** if it is zero or

$$p(\zeta_1, \dots, \zeta_r) \neq 0 \quad \text{whenever} \quad \operatorname{Im} \zeta_i > 0 \quad (1 \leq i \leq r).$$

It is **homogeneous of degree** r if each monomial with a nonzero coefficient has total degree r. For a homogeneous polynomial of degree r in r variables with nonnegative coefficients, define its capacity by

$$\operatorname{Cap}(p) = \inf_{z_1, \dots, z_r > 0} \frac{p(z_1, \dots, z_r)}{z_1 \cdots z_r}. \tag{2.1}$$

The notation $[z_1 \cdots z_r]p$ means the coefficient of that monomial. For integers $r \geq 2$, put

$$g(r) = \left(\frac{r-1}{r}\right)^{r-1}, \qquad g(1) = 1, \qquad \gamma_r = \frac{r!}{r^r}. \tag{2.2}$$

We use the following consequence of Gurvits's theorem [2, Theorem 2.4 and Corollary 2.5]. Gurvits uses the condition $\operatorname{Re} z_i > 0$; for a homogeneous polynomial of degree r, $p(iz) = i^r p(z)$, so that convention is equivalent to the upper-half-plane convention used here. We include the capacity induction and the stability facts it requires.

Proposition 2.1 (Gurvits)**.** *If p is real stable, homogeneous of degree r in r variables, and has nonnegative coefficients, then*

$$[z_1 \cdots z_r]p \geq \gamma_r \operatorname{Cap}(p). \tag{2.3}$$

Proof. Differentiation and specialization of a variable at a real number preserve real stability, allowing the zero polynomial. For specialization, approach the real value from the upper half-plane and apply Hurwitz's theorem on locally uniform limits of nonvanishing holomorphic functions.

For differentiation, suppose the degree in the last variable is $d \geq 1$. Its leading coefficient, as a polynomial in the other variables, does not vanish when those variables have positive imaginary parts: it is the nonzero locally uniform limit of $p(\zeta_1, \dots, \zeta_{r-1}, it)/(it)^d$ as $t \to \infty$. For each such choice of the other variables, the resulting polynomial in the last variable has degree d and all its roots in the closed lower half-plane. Gauss–Lucas gives the same root location for its derivative. Thus the partial derivative is stable. If $d = 0$, the partial derivative is zero.

Let h be a nonzero one-variable polynomial of degree at most r, with nonnegative coefficients and all roots real and nonpositive. We first establish

$$h'(0) \geq g(r) \inf_{t>0} \frac{h(t)}{t}. \tag{2.4}$$

When $h(0) > 0$, write $h(t) = h(0) \prod_{i=1}^{r} (1 + c_i t)$, with $c_i \geq 0$, padding with zeros if necessary. If $u = \sum_i c_i > 0$ and $r \geq 2$, the arithmetic-geometric mean inequality gives $h(t) \leq h(0)(1 + ut/r)^r$. Evaluating this upper bound divided by t at $t = r/[u(r-1)]$ gives (2.4), since $h'(0) = h(0)u$. If $u = 0$, the infimum is zero. For $r = 1$, the assertion is the identity for a linear polynomial. When $h(0) = 0$, nonnegative coefficients give $\inf_{t>0} h(t)/t = h'(0)$, so (2.4) again holds.

Fix positive $z_1, \dots, z_{r-1}$ and apply (2.4) to $h(t) = p(z_1, \dots, z_{r-1}, t)$. Real specialization makes this slice stable; its real coefficients then force all roots to be real, and nonnegative coefficients exclude positive roots. For

$$q(z_1, \dots, z_{r-1}) = \left.\frac{\partial p}{\partial z_r}\right|_{z_r = 0},$$

the definition of capacity gives $q(z) \geq g(r)\operatorname{Cap}(p)\prod_{i<r} z_i$, and hence $\operatorname{Cap}(q) \geq g(r)\operatorname{Cap}(p)$. The closure facts permit iteration until the coefficient is reached. The product of the factors is $\prod_{s=2}^{r} g(s) = \gamma_r$. If a zero polynomial occurs, the preceding capacity bound forces zero capacity, and the conclusion remains valid. This proves (2.3). □

The next lemma refines the scalar step. For a vector $v = (v_i)$, write $\|v\|_2^2 = \sum_i v_i^2$.

Lemma 2.2. *Let $h(t)$ be a polynomial of degree $n \geq 2$ with positive coefficients and only real negative roots. The function $h(t)/t$ has a unique minimizer $t_* > 0$. Write*

$$h(t) = h(0)\prod_{i=1}^{n}(1 + t/\lambda_i), \qquad p_i = \frac{t_*}{t_* + \lambda_i}, \qquad d = \sum_i p_i^2 - \frac{1}{n}, \tag{2.5}$$

where $\lambda_i > 0$. Then $\sum_i p_i = 1$, $d \geq 0$, and

$$h'(0) \geq g(n)\frac{h(t_*)}{t_*}\exp\left[\left(\frac{1}{2} + \frac{1}{3n}\right)d + \frac{d^2}{6}\right]. \tag{2.6}$$

Proof. The derivative of $\log(h(t)/t)$ with respect to $\log t$ is $\sum_i t/(t+\lambda_i) - 1$. It is strictly increasing from -1 to $n-1$, and therefore has exactly one zero. This proves the minimizer claim and $\sum_i p_i = 1$.

Put $C_h = h(t_*)/t_*$. Factorization and weighted arithmetic-geometric mean, with weights p_i, give

$$\frac{h'(0)}{C_h} = \sum_i p_i \prod_{\ell \neq i}(1 - p_\ell) \geq \prod_i (1 - p_i)^{1-p_i}. \tag{2.7}$$

For $0 \leq u < 1$, define $f(u) = (1-u)\log(1-u)$, with the continuous value $f(1) = 0$. Its expansion is

$$f(u) = -u + \sum_{\ell=2}^{\infty} \frac{u^\ell}{\ell(\ell-1)}. \tag{2.8}$$

For each $\ell \geq 2$, convexity gives $\sum_i p_i^\ell \geq n^{1-\ell}$. In addition, Cauchy–Schwarz gives $\sum_i p_i^3 \geq (\sum_i p_i^2)^2$. Thus

$$\begin{aligned} \sum_i f(p_i) - nf(1/n) &\geq \frac{d}{2} + \frac{1}{6}\left(\sum_i p_i^3 - \frac{1}{n^2}\right) \\ &\geq \left(\frac{1}{2} + \frac{1}{3n}\right)d + \frac{d^2}{6}. \end{aligned} \tag{2.9}$$

Since $nf(1/n) = \log g(n)$, (2.7)–(2.9) prove (2.6). □

3 A polynomial for subpermanent sums

We initially assume $A = (a_{ij}) \in \Omega_n$ has strictly positive entries. For $y = (y_1, \ldots, y_n)$, define the elementary symmetric polynomial

$$e_k(y) = \sum_{\substack{S \subseteq [n] \\ |S| = k}} \prod_{i \in S} y_i.$$

Put $m = n - k$, $X = A - J_n$, and define

$$P(z) = \frac{e_k(Az)}{\binom{n}{k}}\left(\frac{z_1 + \cdots + z_n}{n}\right)^m. \tag{3.1}$$

Friedland and Gurvits used the degree-n polynomial $e_k(Az)((z_1 + \cdots + z_n)/n)^{n-k}$ to derive the subpermanent lower bound [10, Proposition 2.5, proof of Theorem 3.1, and Corollary 3.2]. Our P is their polynomial normalized by $\binom{n}{k}$. The additional m linear factors raise the degree from k to n without changing the normalization at $\mathbf{1}$. They will also control the loss incurred when a slice is optimized.

3.1 Stability, capacity, and normalization

The polynomial P is homogeneous of degree n and has a positive coefficient for every monomial of that degree. To see stability, consider $\prod_{i=1}^{n}(y + (Az)_i)$. If y and all z_j have positive imaginary parts, so does every linear factor. Differentiating $n - k$ times in y and then setting $y = 0$ gives $(n - k)!\, e_k(Az)$. Proposition 2.1's closure facts therefore show that $e_k(Az)$ is stable. The same differentiation argument with independent arguments shows that e_k itself is stable. The uniform linear factor in (3.1) is stable too.

For $z > 0$, arithmetic-geometric mean over the $\binom{n}{k}$ subset products gives

$$\frac{e_k(Az)}{\binom{n}{k}} \geq \left(\prod_i (Az)_i\right)^{k/n} \geq \left(\prod_j z_j\right)^{k/n}. \tag{3.2}$$

In the second inequality, apply weighted arithmetic-geometric mean to each row, then use the column sums. Applying ordinary arithmetic-geometric mean to the uniform factor proves $P(z) \geq \prod_j z_j$. Write ∇P for the vector of first partial derivatives, and $\nabla^2 \log P$ for the Hessian, the matrix of second partial derivatives of $\log P$. Since $P(\mathbf{1}) = 1$, we have

$$\mathrm{Cap}(P) = 1, \qquad P(\mathbf{1}) = 1, \qquad \nabla P(\mathbf{1}) = \mathbf{1}. \tag{3.3}$$

For the gradient identity, each derivative of e_k at $\mathbf{1}$ is $\binom{n-1}{k-1}$; composition with A uses its column sums, giving k/n in each coordinate. The uniform factor contributes m/n.

Expanding the coefficient of $z_1 \cdots z_n$ gives

$$[z_1 \cdots z_n]P = \frac{m!}{n^m \binom{n}{k}} \sigma_k(A) = \gamma_n \frac{\sigma_k(A)}{S_{n,k}}. \tag{3.4}$$

Indeed, $e_k(Az)$ contributes each permanent with k distinct columns, and the other m distinct columns must each occur once in the uniform factor. Their coefficient is $m!/n^m$.

Write I_n for the $n \times n$ identity matrix. We also need the second derivatives. The Hessian of $\log e_k$ at $\mathbf{1}$ is

$$-\alpha_{n,k} I_n + \left(\alpha_{n,k} - \frac{k^2}{n^2}\right) \mathbf{1}\mathbf{1}^{\mathsf{T}},$$

because its first derivatives, divided by $e_k(\mathbf{1})$, equal k/n, its pure second derivatives vanish, and its mixed second derivatives, divided by $e_k(\mathbf{1})$, equal $\alpha_{n,k}$. Using $A^{\mathsf{T}}A = J_n + X^{\mathsf{T}}X$ and including the uniform factor yields

$$\nabla^2 \log P(\mathbf{1}) = -J_n - \alpha_{n,k} X^{\mathsf{T}} X. \tag{3.5}$$

3.2 Optimized slices

Fix a column index j and positive coordinates $x = (x_h)_{h \neq j}$. Let

$$h_x(t) = P(z), \qquad z_j = t, \quad z_h = x_h \ (h \neq j).$$

The slice has degree n, positive coefficients, and only real negative roots. Write these roots as $-\lambda_i(x)$, and let $t_*(x)$ be its minimizer from Lemma 2.2. Define

$$\begin{aligned} p_i(x) &= \frac{t_*(x)}{t_*(x) + \lambda_i(x)}, \\ d_j(x) &= \sum_i p_i(x)^2 - \frac{1}{n}, \\ F_j(x) &= \log \frac{h_x(t_*(x))}{t_*(x) \prod_{h \neq j} x_h}, \\ M_j &= \inf_{x>0} \left(F_j(x) + \frac{1}{2} d_j(x) \right). \end{aligned} \tag{3.6}$$

The vector $p(x)$ is a probability vector, meaning that its coordinates are nonnegative and sum to one. Also $F_j(x) \geq 0$ by (3.3). When every $x_h = 1$, (3.3) makes $t_* = 1$. Write $p^0 = p(\mathbf{1})$ and $d_0 = d_j(\mathbf{1})$, suppressing their dependence on the fixed column j. Differentiating the logarithm of the factorization of h_x and using (3.5) gives

$$d_j(\mathbf{1}) = \|p^0 - \mathbf{1}/n\|_2^2 = \alpha_{n,k} \|X_{:j}\|_2^2, \qquad \sum_j d_j(\mathbf{1}) = \Delta_{n,k}(A). \tag{3.7}$$

The notation $X_{:j}$ denotes column j. Any probability vector satisfies $\sum_i (p_i^0)^2 \leq 1$, so $0 \leq d_0 \leq 1 - 1/n$.

For $q_j = \partial P / \partial z_j |_{z_j = 0}$, Lemma 2.2 and (3.6) imply, pointwise in x,

$$q_j(x) \geq g(n) \left(\prod_{h \neq j} x_h \right) \exp\left(F_j(x) + \frac{1}{2} d_j(x) \right).$$

Thus $\mathrm{Cap}(q_j) \geq g(n) e^{M_j}$. The polynomial q_j is stable, homogeneous of degree $n-1$ in $n-1$ variables, and has nonnegative coefficients. Proposition 2.1 applies, and $\gamma_{n-1} g(n) = \gamma_n$. Its coefficient is the one in (3.4). Consequently

$$\frac{\sigma_k(A)}{S_{n,k}} \geq \exp\left(\max_{1 \leq j \leq n} M_j \right). \tag{3.8}$$

Our task is now to bound the infimum M_j in terms of the value d_0 at $\mathbf{1}$. Evaluating (3.6) at $\mathbf{1}$ would give an upper bound on M_j; a lower bound requires control at every positive x.

3.3 Movement of the roots

Assume $k < n$, so $m \geq 1$ and $n \geq 3$. For the fixed column j, put

$$a_i = a_{ij}, \qquad b_i = \sum_{h \neq j} a_{ih} x_h, \qquad b_i^0 = 1 - a_i,$$

and

$$u = \min_{h \neq j} x_h, \qquad v = \max_{h \neq j} x_h, \qquad R = \log(v/u). \tag{3.9}$$

Then $ub^0 \leq b \leq vb^0$ coordinatewise.

Lemma 3.1. *Let $a, b \in (0, \infty)^n$. Viewed as a polynomial in t, $e_k(b + at)$ has k real roots, counted with multiplicity. List them in nondecreasing order. Each listed root is nonincreasing in every coordinate of b, and multiplying b by $c > 0$ multiplies every root by c.*

Proof. Let $\mathbf{e}_i$ denote the i-th standard basis vector. Fix b and increase one coordinate b_i by s. Since e_k has degree at most one in each coordinate, the resulting polynomial has the form $H(t) + sL(t)$, with $\deg H = k$ and $\deg L \le k-1$. It is stable in the two variables t, s, because $e_k(b + at + s\mathbf{e}_i)$ is stable: for positive imaginary parts of t, s, every argument of e_k has positive imaginary part.

For $\operatorname{Im} t > 0$, $H(t) \neq 0$. Because H has real coefficients, its nonreal zeros occur in conjugate pairs. Since H has no zero in the open upper half-plane, all its zeros are real. If $L(t) \neq 0$, the root $s = -H(t)/L(t)$ cannot lie in the upper half-plane. It follows that $\operatorname{Im}(L(t)/H(t)) \le 0$, also when $L(t) = 0$. The real rational function L/H tends to zero at infinity. After cancellation, its poles are real and simple, and its residues are nonnegative. To verify this last assertion directly, expand at a real pole: a pole of order at least two changes the sign of the imaginary part on small upper semicircles, whereas a simple pole $\rho/(t-r)$ has nonpositive imaginary part there precisely when $\rho \ge 0$.

Let the remaining poles be r_ℓ, with residues $\rho_\ell > 0$, and put $w_\ell = \sqrt{\rho_\ell}$. The determinant identity

$$\det\Big(tI - \operatorname{diag}(r_\ell) + sww^{\mathsf{T}}\Big) = \prod_\ell (t - r_\ell)\left(1 + s\sum_\ell \frac{\rho_\ell}{t - r_\ell}\right) \tag{3.10}$$

shows that the moving roots are the eigenvalues of $\operatorname{diag}(r_\ell) - sww^{\mathsf{T}}$. The cancelled factors give constant roots. The variational characterization of ordered eigenvalues shows that subtracting sww^{T} makes every ordered eigenvalue nonincreasing for $s \ge 0$, since $y^{\mathsf{T}}ww^{\mathsf{T}}y = (w^{\mathsf{T}}y)^2 \ge 0$. Adding the constant roots preserves this ordering conclusion, including multiple roots. Finally, $e_k(cb + at) = c^k e_k(b + a(t/c))$ proves the scaling assertion. □

Pair the k roots of the e_k factor with their ordered counterparts at $x = \mathbf{1}$. Lemma 3.1 and the coordinate bounds on b give

$$u\lambda_i^0 \le \lambda_i(x) \le v\lambda_i^0. \tag{3.11}$$

Each of the m uniform factors has root magnitude $\sum_{h \neq j} x_h$, compared with $n-1$ at $\mathbf{1}$, so (3.11) also holds for those roots. We pair these factors separately; no ordering of the combined list is needed.

Abbreviate $p = p(x)$, $d = d_j(x)$, $F = F_j(x)$ and $t = t_*(x)$. The quantities

$$\delta_i = \log \frac{p_i^0/(1-p_i^0)}{p_i/(1-p_i)} = \log \frac{\lambda_i(x)}{t\lambda_i^0}$$

lie in an interval of length R. Since p and p^0 both sum to one, $\min_i \delta_i \le 0 \le \max_i \delta_i$. The identity $p_i^0 = e^{\delta_i} p_i/(1 - p_i + e^{\delta_i} p_i)$ shows that $\log(p_i^0/p_i)$ lies between 0 and δ_i, regardless of its sign. Thus, with $r_i = p_i^0/p_i$,

$$\frac{\max_i r_i}{\min_i r_i} \le e^R, \qquad \sum_i p_i r_i = 1. \tag{3.12}$$

For completeness, if $\ell = \min_i r_i$, $w = \max_i r_i$, and $q = w/\ell$, the inequality $(r_i - \ell)(w - r_i) \ge 0$ yields

$$\sum_i p_i (r_i - 1)^2 \le (w-1)(1-\ell) \le \frac{(q-1)^2}{4q} \le \sinh^2(R/2). \tag{3.13}$$

The middle inequality maximizes $(q\ell - 1)(1 - \ell)$ over ℓ; the last uses $q \le e^R$. Here $\sinh t = (e^t - e^{-t})/2$. Let $p_{\max} = \max_i p_i$. It follows that

$$\|p^0 - p\|_2^2 \le p_{\max} \sinh^2(R/2), \qquad p_{\max} \le \frac{1}{n} + \sqrt{(1 - 1/n)d}. \tag{3.14}$$

For the second bound, if the largest deviation from $1/n$ is a, the remaining deviations sum to $-a$ and their squares sum to at least $a^2/(n-1)$.

4 The lower bound

The root comparison in Section 3 controls dispersion in terms of the range R of the other coordinates. The uniform factors now bound that range by the capacity cost F. Combining the two estimates gives a bound that holds before taking the infimum in (3.6).

4.1 The cost of changing the other coordinates

Continue with $2 \le k < n$ and the notation of Section 3. Define

$$G(z) = n \log \left(\frac{\sum_i z_i}{n} \right) - \sum_i \log z_i.$$

Keeping the uniform factor in (3.2), rather than replacing it by its geometric mean, gives

$$\frac{P(z)}{\prod_i z_i} \ge \left[\frac{(\sum_i z_i/n)^n}{\prod_i z_i} \right]^{m/n}.$$

At $z_j = t_*(x)$ this says

$$F \ge \frac{m}{n} G(z). \tag{4.1}$$

We claim that $G(z) \ge 2 \log \cosh(R/2)$, where $\cosh t = (e^t + e^{-t})/2$. Choose two of the coordinates x_h having ratio $r = v/u$ and scale all coordinates so that these two have arithmetic mean one. Their product is $4r/(1+r)^2$. Let $w > 0$ be the arithmetic mean of the remaining $n-2$ coordinates of the scaled z, including the coordinate indexed by j (which was $t_*(x)$ before scaling). Their product is at most w^{n-2}. Weighted arithmetic-geometric mean gives

$$\left(\frac{2 + (n-2)w}{n} \right)^n \ge w^{n-2}.$$

Since G is unchanged by a common scaling, we obtain

$$G(z) \ge \log \frac{(1+r)^2}{4r} = 2 \log \cosh(R/2). \tag{4.2}$$

This argument also covers $R = 0$, by choosing any two of the other coordinates. Consequently $\sinh^2(R/2) \le e^{nF/m} - 1$. The triangle inequality and (3.14) now yield

$$\sqrt{d_0} \le \sqrt{d} + \sqrt{p_{\max}(e^{nF/m} - 1)}. \tag{4.3}$$

The estimate allows d to vanish away from $\mathbf{1}$; the cost F then accounts for the displacement.

4.2 A pointwise capacity bound

Proposition 4.1. *For every positive $A \in \Omega_n$, every $2 \le k < n$, every column j, and every $x > 0$,*

$$F_j(x) + \frac{1}{2} d_j(x) \ge \frac{7}{12n} d_j(\mathbf{1}). \tag{4.4}$$

Proof. If $d_0 = 0$, the conclusion follows from $F, d \ge 0$. Suppose $d_0 > 0$ and, seeking a contradiction, assume $F + d/2 \le B d_0/n$, where $B = 7/12$. Set

$$s = \sqrt{d/d_0}, \qquad \eta = F/d_0.$$

Then

$$\eta + \frac{s^2}{2} \le \frac{B}{n}, \qquad 0 \le s \le \sqrt{\frac{7}{6n}} < 1, \qquad 0 \le \frac{nF}{m} \le B.$$

For $0 \le y \le 7/12$, the exponential series gives

$$e^y - 1 \le y + \frac{y^2}{2(1 - y/3)} \le \frac{11}{8} y. \tag{4.5}$$

Indeed, $\ell! \ge 2 \cdot 3^{\ell-2}$ for $\ell \ge 2$, and the middle expression divided by y is at most $79/58 < 11/8$. By (3.14) and $d_0 \le 1 - 1/n$, $np_{\max} \le 1 + (n-1)s$. Since $m \ge 1$, (4.5) and the preceding bounds give

$$\begin{aligned} \frac{p_{\max}(e^{nF/m} - 1)}{d_0} &\le \frac{11}{8}(np_{\max})\eta \\ &\le \frac{11}{8}(1 + (n-1)s)\left(\frac{B}{n} - \frac{s^2}{2}\right). \end{aligned}$$

Dividing (4.3) by $\sqrt{d_0}$ and using this bound gives

$$1 \le s + \sqrt{\frac{11}{8}(1 + (n-1)s)\left(\frac{7}{12n} - \frac{s^2}{2}\right)}. \tag{4.6}$$

For fixed $0 \le s < 1$, the expression under the square root, before its factor $11/8$, equals

$$Bs + \frac{B(1-s)}{n} - \frac{s^2(1-s)}{2} - \frac{ns^3}{2}.$$

It decreases as n increases and hence is at most its value at $n = 3$. However,

$$\begin{aligned} \Phi(s) &= (1-s)^2 - \frac{11}{8}(1 + 2s)\left(\frac{7}{36} - \frac{s^2}{2}\right) \\ &= \frac{211}{288} - \frac{365}{144}s + \frac{27}{16}s^2 + \frac{11}{8}s^3 > 0 \qquad (s \ge 0). \end{aligned} \tag{4.7}$$

To check positivity entirely with rational constants, put $w = s - 1/2$. Since $w \ge -1/2$,

$$\begin{aligned} \Phi(s) &= \frac{17}{288} + \frac{53}{288}w + \left(\frac{15}{4} + \frac{11}{8}w\right)w^2 \\ &\ge \frac{17}{288} + \frac{53}{288}w + \frac{49}{16}w^2 \\ &\ge \frac{57167}{1016064} > 0. \end{aligned} \tag{4.8}$$

The last inequality minimizes the quadratic. In the range of s under consideration, $1 - s > 0$ and the square root in (4.6) is defined. Equations (4.7)–(4.8) make that square root strictly less than $1 - s$, contradicting (4.6). This proves (4.4). □

Taking the infimum in (4.4) gives $M_j \ge 7d_j(\mathbf{1})/(12n)$. By (3.7), some column satisfies $d_j(\mathbf{1}) \ge \Delta_{n,k}(A)/n$. Therefore (3.8) proves

$$\frac{\sigma_k(A)}{S_{n,k}} \ge \exp\left(\frac{7}{12n^2}\Delta_{n,k}(A)\right) \qquad (2 \le k < n). \tag{4.9}$$

4.3 Permanent order and the boundary

When $k = n$, the uniform factor is absent. In this case $P(z) = \prod_i (Az)_i$ is a product of row forms, and a direct one-column argument is available. This estimate is implicit in the proof of Laurent and Schrijver [3, proof of Corollary 1d, equation (39)].

Fix j, put $a_i = a_{ij}$ and $L_i(x) = \sum_{h \neq j} a_{ih} x_h$. Differentiating at $z_j = 0$ gives

$$q_j(x) = \sum_i a_i \prod_{\ell \neq i} L_\ell(x) \geq \prod_i L_i(x)^{1-a_i}.$$

For each row, its remaining sum is $1-a_i$, so a second weighted arithmetic-geometric mean inequality gives

$$L_i(x)^{1-a_i} \geq (1-a_i)^{1-a_i} \prod_{h \neq j} x_h^{a_{ih}}.$$

Multiplying and using the remaining column sums proves $\operatorname{Cap}(q_j) \geq \prod_i (1-a_i)^{1-a_i}$. Proposition 2.1 and $\gamma_{n-1} g(n) = \gamma_n$ imply

$$\log \frac{\operatorname{per} A}{\gamma_n} \geq \sum_i f(a_i) - n f(1/n) \geq \frac{1}{2} \sum_i (a_i - 1/n)^2. \tag{4.10}$$

The final inequality follows from (2.8), since $\sum_i a_i = 1$ and every power-sum difference of degree at least two is nonnegative; the quadratic difference is $\sum_i (a_i - 1/n)^2$. Choosing a column with squared deviation at least $\|A - J_n\|_F^2/n$ gives

$$\frac{\operatorname{per} A}{\gamma_n} \geq \exp\left(\frac{1}{2n} \|A - J_n\|_F^2 \right). \tag{4.11}$$

Here $1/(2n) \geq 7/(12n^2)$ for $n \geq 2$, while $S_{n,n} = \gamma_n$ and $\alpha_{n,n} = 1$. Thus (4.11) proves the lower estimate in Theorem 1.2 at $k = n$.

For an arbitrary nonnegative $A \in \Omega_n$, apply the lower estimate to $A_\varepsilon = (1-\varepsilon)A + \varepsilon J_n$, $0 < \varepsilon < 1$. Its entries are positive, and $A_\varepsilon - J_n = (1-\varepsilon)(A - J_n)$. Letting $\varepsilon \downarrow 0$ proves the same non-strict estimate for A by continuity of the subpermanent sums and the norm. The strict transform comparison will follow from the positive coefficient in (1.7), so it requires no passage of a strict inequality through a limit.

5 A tensor estimate for centered subpermanents

A real $n \times n$ matrix has **zero margins** if every row and column sum is zero. Write $\|X\|_{2\to 2} = \sup_{v \neq 0} \|Xv\|_2/\|v\|_2$, $Q = I_n - J_n$, and $q_{n,r} = \sigma_r(Q)$. The centered subpermanent estimate is [4, Lemma 3.1], and $q_{n,r} \geq 0$ follows from [4, equation (3.5)].

Lemma 5.1. *For $n \geq 2$, if X has zero margins and $\|X\|_{2\to 2} \leq 1$, then for $2 \leq r \leq n$,*

$$q_{n,r} \geq 0, \qquad |\sigma_r(X)| \leq \frac{q_{n,r}}{n-1} \|X\|_F^2. \tag{5.1}$$

For $A \in \Omega_n$ and $v \in \mathbb{R}^n$, weighted Cauchy–Schwarz in each row gives

$$\sum_i \left(\sum_j a_{ij} v_j \right)^2 \leq \sum_{i,j} a_{ij} v_j^2 = \sum_j v_j^2. \tag{5.2}$$

Therefore $\|A\|_{2\to 2} \leq 1$. The row and column sums give $A - J_n = QAQ$, so $X = A - J_n$ has zero margins and $\|X\|_{2\to 2} \leq 1$. Lemma 5.1 applies to every such A, including those with zero entries.

6 The upper bound and completion of the proof

6.1 Expanding the transform

For $0 \le r \le k$, put

$$c_{n,k,r} = (k-r)!\binom{n-r}{k-r}^2 n^{-(k-r)}.$$

For every real t and every $n \times n$ matrix X,

$$\sigma_k(J_n + tX) = \sum_{r=0}^{k} c_{n,k,r} t^r \sigma_r(X). \tag{6.1}$$

This identity is also a specialization of Kopotun's Lemma B [1, equation (26)]; we include the counting argument for completeness. To verify the multiplicity, choose the r entries from X in a partial matching, meaning a collection of positions with distinct row indices and distinct column indices. Its rows and columns can be extended to k rows and columns in $\binom{n-r}{k-r}^2$ ways, and the remaining $k-r$ rows can be matched in $(k-r)!$ ways. Each entry supplied by J_n contributes $1/n$.

Set $\rho = 1/(n+1)$ and $X = A - J_n$. Since $\sigma_1(X) = 0$ and $c_{n,k,0} = S_{n,k}$, Lemma 5.1 gives

$$\sigma_k(\tau_n(A)) \le S_{n,k} + U_{n,k}\|X\|_F^2, \qquad U_{n,k} = \frac{1}{n-1}\sum_{r=2}^{k} c_{n,k,r}\rho^r q_{n,r}. \tag{6.2}$$

All summands defining $U_{n,k}$ are nonnegative. Define the falling factorial by $(a)_r = a(a-1)\cdots(a-r+1)$ for $r \ge 1$, and $(a)_0 = 1$. Direct cancellation gives

$$\frac{c_{n,k,r}}{S_{n,k}} = \frac{n^r (k)_r}{(n)_r^2}. \tag{6.3}$$

For each $r \ge 2$, division by $\alpha_{n,k} = (k)_2/(n)_2$ leaves a positive factor independent of k times $(k-2)_{r-2}$. This is nondecreasing for integers $k \ge r$. Increasing k also introduces only nonnegative summands in (6.2). Consequently

$$\frac{U_{n,k}}{S_{n,k}\alpha_{n,k}} \le \frac{U_{n,n}}{S_{n,n}}. \tag{6.4}$$

At $k = n$, (6.1) with $X = Q$ identifies the right-hand side as

$$\frac{1}{n-1}\left(\frac{\operatorname{per}(J_n + \rho Q)}{\gamma_n} - 1\right). \tag{6.5}$$

Now $J_n + \rho Q = (I_n + \mathbf{1}\mathbf{1}^{\mathsf{T}})/(n+1)$. Expanding the permanent according to the identity entries selected gives

$$\operatorname{per}(I_n + \mathbf{1}\mathbf{1}^{\mathsf{T}}) = \sum_{r=0}^{n}\binom{n}{r}(n-r)! = n!\sum_{r=0}^{n}\frac{1}{r!}. \tag{6.6}$$

Thus (6.5) equals b_n. Combining (6.2)–(6.6) proves the upper estimate in Theorem 1.2. It also shows $b_n > 0$: the $r = 2$ summand of $U_{n,n}$ in (6.2) is positive, since $q_{n,2} = (n-1)/2$. This last identity follows from $2\sigma_2(X) = \sum_{i\ne j,\ h\ne \ell} x_{ih}x_{j\ell} = \sum_{i,h} x_{ih}^2$ for any zero-margin X: inclusion-exclusion eliminates the total sum and the terms with equal rows or equal columns. Apply this to Q, whose squared Frobenius norm is $n-1$.

6.2 Separating the constants

It remains to prove $n^2 b_n < 7/12$. Direct evaluation for $n = 2, 3, 4$ gives, respectively,

$$n^2 b_n = \frac{4}{9}, \qquad \frac{9}{16}, \qquad \frac{656}{1125}, \tag{6.7}$$

each strictly less than $7/12$.

For $n \geq 5$, the alternating expansion of $\log(1 + 1/n)$ and $\sum_{r=0}^{n} 1/r! < e$ give

$$\left(\frac{n}{n+1}\right)^n \sum_{r=0}^{n} \frac{1}{r!} < e^{y_n}, \qquad y_n = \frac{1}{2n} - \frac{1}{3n^2} + \frac{1}{4n^3}. \tag{6.8}$$

Indeed, $1 - n\log(1 + 1/n) < y_n$. Also $0 < y_n < 1$. Since $r! \geq 2^{r-1}$ for $r \geq 1$, summing the exponential series gives

$$e^y - 1 \leq \frac{y}{1 - y/2} \qquad (0 \leq y < 2). \tag{6.9}$$

It suffices to show $y_n/(1-y_n/2) < 7(n-1)/(12n^2)$. Clearing positive denominators, this is equivalent to

$$24n^4 - 114n^3 - 2n^2 - 49n + 21 > 0. \tag{6.10}$$

Writing $n = 5 + t$, the polynomial in (6.10) is

$$476 + 3381t + 1888t^2 + 366t^3 + 24t^4,$$

which is positive for $t \geq 0$. Equations (6.7)–(6.10) prove $b_n < C_n$ for every $n \geq 2$. Together with Section 4 and the upper bound, this completes Theorem 1.2.

6.3 Equality and the quantitative gap

Proof of Theorem 1.1. Subtract the estimates of Theorem 1.2 and use $e^u \geq 1 + u$ for $u \geq 0$ to obtain (1.7). Since $S_{n,k} > 0$, $\alpha_{n,k} > 0$, and $C_n - b_n > 0$, that inequality is strict whenever $A \neq J_n$. At $A = J_n$ the transform fixes A, giving equality. Thus Theorem 1.1 holds in every stated order, including all support boundaries. □

Acknowledgments

The proof of this conjecture was carried out by GPT-5.6-sol, GPT-6 Astra, Claude Fable 5, and Claude Fable 5.1, under the guidance of the author. The author has reviewed the resulting proof arguments. Responsibility for the final text rests with the author.

References

[1] K. A. Kopotun, "A note on the convexity of the sum of subpermanents," *Linear Algebra and its Applications* **245** (1996), 157–169. doi:10.1016/0024-3795(94)00228-2.

[2] L. Gurvits, "Van der Waerden/Schrijver-Valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: One theorem for all," *Electronic Journal of Combinatorics* **15** (2008), Research Paper 66. arXiv:0711.3496.

[3] M. Laurent and A. Schrijver, "On Leonid Gurvits's proof for permanents," *American Mathematical Monthly* **117** (2010), 903–911. doi:10.4169/000298910X523380.

[4] Y. Lavi, "The Marcus-Minc Transform Inequality," preprint (2026). arXiv:2609.29262v1.

[5] Y. Lavi, "The Wang Transform Inequality," preprint (2026). arXiv:2609.31794v1.

[6] E. T. H. Wang, "On a conjecture of M. Marcus and H. Minc," *Linear and Multilinear Algebra* **5** (1977), 145–148. doi:10.1080/03081087708817189.

[7] D. K. Chang, "Minimum and maximum permanents of certain doubly stochastic matrices," *Linear and Multilinear Algebra* **24** (1988), 39–44. doi:10.1080/03081088808817895.

[8] I. M. Wanless, "The Holens–Đoković conjecture on permanents fails!" *Linear Algebra and its Applications* **286** (1999), 273–285. doi:10.1016/S0024-3795(98)10177-5.

[9] S. Friedland, "A proof of a generalized van der Waerden conjecture on permanents," *Linear and Multilinear Algebra* **11** (1982), 107–120. doi:10.1080/03081088208817436.

[10] S. Friedland and L. Gurvits, "Generalized Friedland–Tverberg inequality: applications and extensions," arXiv:math/0603410v2 (2006). Shorter published version: "Lower Bounds for Partial Matchings in Regular Bipartite Graphs and Applications to the Monomer–Dimer Entropy," *Combinatorics, Probability and Computing* **17** (2008), 347–361. doi:10.1017/S0963548307008747.