hexagon:2610.00731v1

# A SHARP BRÉGMAN–MINC BOUND FOR PERMANENTS

SUVRIT SRA*

ABSTRACT. We prove a sharp Brégman–Minc bound for permanents of nonnegative matrices that was conjectured by Samorodnitsky and Gurvits. Specifically, for $\phi(x) = \Gamma(1 + 1/x)^{-x}$, with $\phi(0) = 0$, applied entrywise, we show that every nonnegative row-stochastic square matrix $A$ satisfies $\operatorname{per} \phi[A] \le 1$, which is the Brégman–Minc inequality when each row of $A$ is uniform on its support. The crux of our proof is a scalar inequality on the simplex, which we prove using a carefully chosen yet elementary integral representation. We note brief applications to the entropy of a random permutation and to sharp bounds for rectangular permanents and matchings.

## 1. INTRODUCTION

The permanent of an $n \times n$ matrix $B = (b_{ij})$ is

$$\operatorname{per} B = \sum\nolimits_{\sigma \in \mathfrak{S}_n} \prod\nolimits_{i=1}^{n} b_{i,\sigma(i)},$$

where $\mathfrak{S}_n$ denotes the permutations of $[n] := \{1, \dots, n\}$. If $B \in \{0,1\}^{n \times n}$ with positive row sums $r_1, \dots, r_n$, the Brégman–Minc inequality (Brégman, 1973) states that

$$\operatorname{per} B \le \prod\nolimits_{i=1}^{n} (r_i!)^{1/r_i}. \tag{1.1}$$

Seeking to extend (1.1) to arbitrary nonnegative matrices, Samorodnitsky (2008) asked which function of a row plays the role of $(r!)^{1/r}$. He proposed the function

$$\phi(x) := \Gamma(1 + 1/x)^{-x} \quad (x > 0), \qquad \phi(0) := 0,$$

as a natural extension of $\phi(1/r) = (r!)^{-1/r}$ and conjectured (Samorodnitsky, 2008, Conjecture 1.6) that $\operatorname{per} \phi[A] \le 1$ (applying $\phi$ entrywise) for every *row-stochastic* $A$. The conjecture also appears in Gurvits and Samorodnitsky (2014, Conjecture 2.3). We prove it, including its equality cases.

**Theorem 1.1** (Samorodnitsky–Gurvits inequality)**.** *Let $A \in \mathbb{R}_+^{n \times n}$ be row-stochastic. Then,*

$$\operatorname{per} \phi[A] \le 1.$$

*Equality holds if and only if independent permutations of the rows and columns transform $A$ into*

$$\tfrac{J_{r_1}}{r_1} \oplus \cdots \oplus \tfrac{J_{r_k}}{r_k}, \qquad r_1 + \cdots + r_k = n,$$

*where $J_r$ is the $r \times r$ all-ones matrix.*

There is a valuable reformulation of Theorem 1.1 using Orlicz norms. Consider the function $\psi = \phi^{-1}$ on $[0,1]$ (increasing and convex by Lemma 2.2); using $\psi$ one defines the *Orlicz norm*:

$$\|b\|_\psi := \min\{r \ge \max_j |b_j| : \sum\nolimits_j \psi(|b_j|/r) \le 1\}, \tag{1.2}$$

2020 *Mathematics Subject Classification.* Primary 15A15; Secondary 26D15, 46B20, 94A17.

*Key words and phrases.* permanent, Gamma function, Orlicz norm, Brégman–Minc inequality, entropy bound.

* TU Munich, Dept. of Mathematics, School of CIT; Garching, Germany. Email: s.sra@tum.de.

First public release: 31 October 2026, as v1 in the author's preprint repository, https://github.com/suvrit/preprints/tree/main/preprints/gurvits; every version is kept there with its timestamp.

where we set $\|b\|_\psi = 0$ for $b = 0$. Normalizing each row of a nonnegative matrix in this norm gives an equivalent bound that was conjectured by Samorodnitsky (2008, §1.2), and also stated by Gurvits and Samorodnitsky (2014, §2.2) within a general family of Orlicz permanent bounds.

**Corollary 1.2** (Orlicz permanent inequality)**.** *For every $B \in \mathbb{R}_+^{n\times n}$, with rows $b_i$,*

$$\operatorname{per} B \le \prod_{i=1}^n \|b_i\|_\psi. \tag{1.3}$$

*If all rows are nonzero, equality holds precisely when independent row and column permutations make $B$ block diagonal with square blocks, each of whose rows is a positive constant across that block.*

The Orlicz norm is the function of a row that replaces $(r!)^{1/r}$: for $b = c\mathbf{1}_S$ with $|S| = r \ge 1$,

$$\|b\|_\psi = |c|(r!)^{1/r}, \tag{1.4}$$

so on zero-one matrices Corollary 1.2 is the Brégman–Minc inequality. Soules (2005, Theorem 1) extends that inequality to nonnegative matrices by writing each row as a nonnegative combination of nested zero-one rows and taking the same combination of their Brégman–Minc constants. By the triangle inequality and (1.4), each row factor in Soules's bound is at least the Orlicz norm of the row, so Corollary 1.2 refines Soules's bound row by row.

Other row norms lead to different sharp problems. For prescribed $\ell_p$ row norms, Bristiel and Caputo (2024, Corollary 1.14 in the arXiv version) proved

$$\operatorname{per} B \le \max\{1, \tfrac{n!}{n^{n/p}}\} \prod_i \|b_i\|_p, \qquad p \ge 1,$$

resolving Samorodnitsky (2008, Conjecture 1.1). There the constant depends on $n$, whereas the bound (1.3) uses the same norm in every dimension and is an equality at every block size in Theorem 1.1. For other Orlicz functions $\psi$, Gurvits and Samorodnitsky (2014, Theorem 4.1) proved (1.3) under three conditions on $\psi$, met for instance by $\psi(x) = 1 - (1-x)a^x$ with $a \approx 1.54$. These results leave Samorodnitsky's conjecture open.

We prove Theorem 1.1 by induction on $n$. Deleting a row and a column and renormalizing the remaining rows leaves a row-stochastic minor, and Hölder's inequality reduces the induction step to a scalar inequality: for each row $a$ of $A$, the weights of these minors sum to at most one. For $n \ge 2$ and $a$ in the relative interior of the simplex $\Delta_n = \{a \in [0,1]^n : \sum_i a_i = 1\}$, the weights $q_i(a)$ and their sum $Q_n(a)$ are

$$q_i(a) = \frac{\phi(a_i)^n}{\prod_{j\ne i} \phi\big(\frac{a_i}{1-a_j}\big)}, \qquad Q_n(a) = \sum_{i=1}^n q_i(a). \tag{1.5}$$

These expressions extend continuously to the closed simplex (Section 3.4).

**Theorem 1.3** (Scalar simplex inequality)**.** *Let $n \ge 2$. For every $a \in \Delta_n$,*

$$Q_n(a) \le 1.$$

*Equality holds if and only if the positive coordinates of $a$ are all equal.*

We prove Theorem 1.3 in Section 3 through a positive integral representation of $-\log\phi$, suggested by the integral representations of subentropy in Sra (2019, Section 3), which build on Jozsa and Mitchison (2015); $Q_n$ itself is not a subentropy. The proof also gives a strict form of Theorem 1.1 that is computable from one-dimensional integrals (Proposition 4.1).

Beyond Brégman–Minc, Corollary 1.2 gives sharp bounds for rectangular permanents and for the total weights of $k$-matchings in weighted bipartite graphs (Theorem 5.1, Corollary 5.2). Theorem 1.1 also has an entropy form, which bounds the entropy of a random permutation by the sum of $\mathcal{H}_\Gamma$ over its row marginals, where $\mathcal{H}_\Gamma$ is the entropy functional dual to the Orlicz norm (Theorem 5.4).

## 2. The Gamma function and the Orlicz norm

A positive integral representation of $-\log\phi$ gives both the concavity needed for the norm (1.2) and the integrals used to prove Theorem 1.3. Put

$$h(x) = -\log\phi(x) = x\log\Gamma(1+1/x), \qquad \omega(t) = t\lfloor 1/t\rfloor \quad (0<t<1).$$

On $(0,1)$ we have $0 < 1-t \le \omega(t) \le 1$, and we set $\omega = 0$ outside $(0,1)$ when necessary.

**Lemma 2.1** (Integral representation and recurrence)**.** *For $x > 0$,*

$$h(x) = -\gamma + \int_0^1 \frac{\omega(t)}{x+t}\,dt, \tag{2.1}$$

*where $\gamma$ is Euler's constant. For $0 < x < 1$,*

$$\phi(x) = x^x \phi\left(\frac{x}{1-x}\right)^{1-x}. \tag{2.2}$$

*Proof.* Taking logarithms in Weierstrass's product for the Gamma function (National Institute of Standards and Technology, 2026, Equation 5.8.2) gives

$$h(x) = -\gamma + \sum\nolimits_{m=1}^{\infty}\left\{\frac{1}{m} - x\log\left(1+\frac{1}{mx}\right)\right\}.$$

The $m$-th summand equals

$$\int_0^{1/m} \frac{t}{x+t}\,dt.$$

All summands are nonnegative, so Tonelli's theorem and $\sum_{m\ge1}\mathbf{1}_{\{t\le1/m\}} = \lfloor 1/t\rfloor$ prove (2.1). The recurrence $\Gamma(1+1/x) = x^{-1}\Gamma(1/x)$ proves (2.2). □

The concavity of $\phi$ is known; Samorodnitsky (2008) attributes it to D. Moews. Lemma 2.2 rederives it from (2.1), since the same representation drives the proof of Theorem 1.3.

**Lemma 2.2** (Monotonicity and concavity)**.** *The function $\phi$ is strictly increasing and strictly concave on $(0,\infty)$, and it satisfies*

$$\phi(0+) = 0, \qquad \phi(1) = 1, \qquad \lim_{x\to\infty}\phi(x) = e^{\gamma}.$$

*In particular, $\phi(x) \ge x$ for $0 \le x \le 1$, strictly for $0 < x < 1$.*

*Proof.* Let $L = \log\phi$. Differentiating (2.1) gives

$$L'(x) = \int_0^1 \frac{\omega(t)}{(x+t)^2}\,dt > 0, \qquad -L''(x) = 2\int_0^1 \frac{\omega(t)}{(x+t)^3}\,dt.$$

After the substitution $u = 1/(x+t)$, the two integrals become

$$L'(x) = \int_0^{\infty} f_x(u)\,du, \qquad -L''(x) = 2\int_0^{\infty} u f_x(u)\,du,$$

where $0 \le f_x \le 1$, and $f_x$ is supported in $[1/(x+1), 1/x]$. Every integrable $0 \le f \le 1$ with finite first moment satisfies, on writing the square as a double integral,

$$\left(\int_0^{\infty} f\right)^2 = 2\int_0^{\infty} f(u)\left(\int_0^u f(v)\,dv\right)du \le 2\int_0^{\infty} u f(u)\,du.$$

The inequality is strict for $f_x$, since it is nonzero with support bounded away from zero. Consequently

$$\frac{\phi''(x)}{\phi(x)} = L''(x) + L'(x)^2 < 0.$$

The lower bound $\omega(t) \ge 1-t$ gives

$$h(x) \ge -\gamma + (1+x)\log(1+1/x) - 1 \longrightarrow +\infty \quad (x\downarrow 0).$$

The remaining endpoint assertions follow from $\Gamma(2) = 1$ and (2.1), and strict concavity together with the endpoint values at 0 and 1 gives the final assertion. □

**Proposition 2.3** (Norm properties)**.** *The expression* (1.2) *defines an absolute, coordinatewise monotone norm on every finite-dimensional real or complex vector space. It is consistent under adjoining zero coordinates. Its unit ball is*

$$\{b : \ |b_j| \le 1 \text{ for all } j, \quad \textstyle\sum_j \psi(|b_j|) \le 1\}.$$

*For a nonzero vector $b$, $\|b\|_\psi \ge \|b\|_\infty$, and* (1.4) *holds.*

*Proof.* The inverse $\psi : [0,1] \to [0,1]$ is continuous, increasing, and convex. The displayed unit ball is therefore convex, balanced, and bounded, and it contains a neighborhood of zero, so its Minkowski functional is a norm. For nonzero $b$ and $r \ge \|b\|_\infty$, the sum in (1.2) is continuous and strictly decreasing, starts at a value at least 1, and tends to zero, which proves the normalization and uniqueness. The remaining claims follow from the unit ball. For $c \neq 0$ and $|S| = r$, the identity $r\psi(|c|/\|c\mathbf{1}_S\|_\psi) = 1$ proves (1.4). For $c = 0$, the formula follows from $\|0\|_\psi = 0$. □

*Remark* 2.4. One may instead define $\psi = \phi^{-1}$ on $[0, e^\gamma)$ and $\psi = +\infty$ on $[e^\gamma, \infty)$. The usual Luxemburg functional then gives the same norm, because a sum of $\psi$-values at most 1 forces each argument to be at most 1. No increasing convex extension past 1 affects the statements here.

## 3. The scalar inequality

The induction of Section 4 closes once every row satisfies $Q_n \le 1$, so we prove Theorem 1.3 first.

3.1. **Reduction to two variables.** Each factor in the denominator of (1.5) compares $h$ at $a_i$ and at $a_i/(1-a_j)$. For $0 < x, b < 1$ we therefore set

$$g_x(b) = \frac{h(x) - h(x/(1-b))}{b}, \qquad R_b(x) = \exp\{g_x(x) - g_x(b)\}. \tag{3.1}$$

Substituting (2.1) gives the positive integral representation

$$g_x(b) = \int_0^1 \frac{\omega(t)x}{(x+t)(x+(1-b)t)}\,dt, \tag{3.2}$$

and the recurrence (2.2) gives

$$(1-x)g_x(x) = -h(x) - \log x. \tag{3.3}$$

In particular, $R_x(x) = 1$.

**Lemma 3.1** (Two-variable comparison)**.** *If $a, b > 0$ and $a + b \le 1$, then*

$$R_b(a) + R_a(b) \le 2, \tag{3.4}$$

*with equality if and only if $a = b$.*

Lemma 3.1 is proved in Section 3.3; we first show that it implies the scalar theorem in the interior of the simplex. For $a$ in the relative interior of $\Delta_n$, (3.1) and (3.3) yield

$$\begin{aligned}\log \frac{q_i(a)}{a_i} &= -h(a_i) - \log a_i - \textstyle\sum_{j\neq i} a_j g_{a_i}(a_j) \\ &= \textstyle\sum_{j=1}^n a_j (g_{a_i}(a_i) - g_{a_i}(a_j)).\end{aligned}$$

Exponentiating gives

$$q_i(a) = a_i \textstyle\prod_{j=1}^n R_{a_j}(a_i)^{a_j}.$$

Weighted AM–GM and Lemma 3.1 imply

$$\begin{aligned} Q_n(a) &\le \sum\nolimits_{i,j} a_i a_j R_{a_j}(a_i) \\ &= \sum\nolimits_i a_i^2 + \sum\nolimits_{i<j} a_i a_j \bigl(R_{a_j}(a_i) + R_{a_i}(a_j)\bigr) \\ &\le \sum\nolimits_i a_i^2 + 2\sum\nolimits_{i<j} a_i a_j = 1. \end{aligned} \tag{3.5}$$

Lemma 3.1 applies because $a_i + a_j \le 1$ for distinct indices, and by its equality case the last inequality is strict if two coordinates of $a$ differ.

3.2. **A common positive measure.** To compare $R_b(a)$ with $R_a(b)$, we write the logarithm of each as an integral against the same positive measure. Subtracting the kernels in (3.2) gives

$$\log R_b(a) = (a-b)K(a,b), \qquad K(a,b) = \int_0^1 \frac{\omega(t)\,at}{(a+t)(a+(1-a)t)(a+(1-b)t)}\,dt. \tag{3.6}$$

In particular $K(a,b) > 0$. Now, fix $0 < a < b$, $a+b \le 1$, and write $d = b-a$, $s = a+b$, and $c = s/3$. The substitution $y = a/t$ in (3.6) gives

$$K(a,b) = \int_0^\infty \frac{a\lfloor y/a \rfloor}{y(y+1)(y+1-a)(y+1-b)}\,dy.$$

We name the denominator and its tail integral:

$$P(y) = y(y+1)(y+1-a)(y+1-b), \qquad F(z) = \int_z^\infty \frac{dy}{P(y)}.$$

The polynomial $P$ is symmetric in $a, b$, and Tonelli's theorem gives

$$K(a,b) = a\sum\nolimits_{m\ge1} F(ma), \qquad K(b,a) = b\sum\nolimits_{m\ge1} F(mb). \tag{3.7}$$

The identity

$$\frac{1}{y+\lambda} = \int_0^\infty e^{-yt}e^{-\lambda t}\,dt \qquad (\lambda \ge 0)$$

shows that $1/P$ is the Laplace transform of a nonnegative density $\rho$, namely the convolution on $[0,\infty)$ of $1, e^{-t}, e^{-(1-a)t}, e^{-(1-b)t}$. Therefore

$$F(z) = \int_0^\infty e^{-zt}\,d\nu(t), \qquad d\nu(t) = \frac{\rho(t)}{t}\,dt.$$

Summing the geometric series in (3.7), we obtain

$$A := K(a,b) = \int_0^\infty k_a(t)\,d\nu(t), \qquad B := K(b,a) = \int_0^\infty k_b(t)\,d\nu(t), \quad k_x(t) = \frac{x}{e^{xt}-1}. \tag{3.8}$$

The same positive measure $\nu$ occurs in both expressions. The original integrals in (3.6) are finite, and every exchange above involves nonnegative integrands.

3.3. **Proof of the two-variable comparison.** By (3.6), $R_b(a) = e^{-dA}$ and $R_a(b) = e^{dB}$, so Lemma 3.1 compares the two exponents $dA$ and $dB$. The inequality $A > B$ alone would not suffice; the next lemma bounds $1/B - 1/A$ from below, using the common measure of (3.8).

**Lemma 3.2.** *With the notation above,*

$$\frac{1}{B} - \frac{1}{A} \ge 2d. \tag{3.9}$$

*Proof.* For fixed $t > 0$,

$$\frac{1}{k_x(t)} = t \int_0^1 e^{xtv}\, dv.$$

Subtracting, and writing the difference of exponentials as an integral over $u \in [0,1]$, we obtain

$$\frac{1}{k_b(t)} - \frac{1}{k_a(t)} = d\, t^2 \int_0^1 \int_0^1 v e^{(a+du)tv}\, du\, dv.$$

The measure $du\, 2v\, dv$ is a probability measure on the unit square, under which $(a+du)v$ has mean

$$(a + d/2)(2/3) = s/3 = c.$$

Jensen's inequality gives

$$\frac{1}{k_b(t)} - \frac{1}{k_a(t)} \geq \frac{d\, t^2}{2} e^{ct}.$$

In particular $k_a > k_b$, so $A > B > 0$.

Cauchy–Schwarz turns this pointwise bound into a bound for $AB/(A-B)$:

$$\begin{aligned} \frac{AB}{A-B} &= B + \frac{B^2}{A-B} \\ &\leq \int_0^\infty \left( k_b + \frac{k_b^2}{k_a - k_b} \right) d\nu = \int_0^\infty \frac{d\nu}{1/k_b - 1/k_a} \\ &\leq \frac{2}{d} \int_0^\infty \frac{e^{-ct}}{t^2}\, d\nu(t) = \frac{J}{d}, \end{aligned} \tag{3.10}$$

where, by Tonelli's theorem and $\int_c^\infty (y-c)^2 e^{-yt}\, dy = 2e^{-ct}/t^3$, the constant $J$ is

$$J = 2 \int_0^\infty \frac{e^{-ct}}{t^2}\, d\nu(t) = \int_c^\infty \frac{(y-c)^2}{y(y+1)(y+1-a)(y+1-b)}\, dy. \tag{3.11}$$

It remains to show that $J \leq 1/2$. Since $a + b = 3c$,

$$(y+1-a)(y+1-b) = (y+1)^2 - 3c(y+1) + ab \geq (y+1)(y+1-3c).$$

Set $y = z + c$. The denominator in (3.11) is at least

$$(z+c)(z+1+c)^2(z+1-2c) \geq z(z+1)^3 \qquad (z \geq 0,\ 0 \leq c \leq 1/3).$$

To verify the second inequality, put $r = z + 1$; the difference of its two sides equals

$$c(r^2(r-3c) + c(3-5c)r + 2c^2(1-c)) \geq 0,$$

which is nonnegative because $r \geq 1$ and $c \leq 1/3$. Hence

$$J \leq \int_0^\infty \frac{z}{(z+1)^3}\, dz = \frac{1}{2}.$$

Combining this bound with (3.10) proves $AB/(A-B) \leq 1/(2d)$, which is (3.9). □

***Proof of Lemma 3.1.*** The case $a = b$ is immediate. By symmetry assume $a < b$, and set $u = dA$ and $v = dB$. Lemma 3.2 implies

$$\frac{1}{v} - \frac{1}{u} \geq 2, \qquad 0 < v \leq \frac{u}{1+2u} < \frac{1}{2}.$$

The elementary exponential inequalities $e^v - 1 \leq v/(1-v)$ for $0 \leq v < 1$ and $1 - e^{-u} \geq u/(1+u)$ for $u \geq 0$ give the chain

$$e^v - 1 \leq \frac{v}{1-v} \leq \frac{u}{1+u} \leq 1 - e^{-u}.$$

Since $R_b(a) = e^{-u}$ and $R_a(b) = e^v$, this chain proves (3.4). If $a \neq b$, then $u, v > 0$, and the elementary exponential inequalities are strict. □

### 3.4. **Boundary continuity and equality.**

***Completion of the proof of Theorem 1.3.*** In the relative interior, the inequality was proved in (3.5). Since $\phi$ is increasing and $a_i/(1-a_j) \geq a_i$, one has

$$0 \leq q_i(a) \leq \phi(a_i).$$

Thus a summand tends to zero when its coordinate tends to zero. For a positive coordinate, each zero coordinate elsewhere supplies a factor $\phi(a_i)$ in its denominator, cancelling one numerator factor. Consequently $Q_n$ restricts to $Q_r$ on a face with $r$ positive coordinates. At a vertex the nonzero-coordinate summand tends to 1, and all others tend to zero. These limits prove continuity and the inequality on the closed simplex.

If two positive coordinates differ, the corresponding term in (3.5) is strict. Conversely, at a vector uniform on $r$ coordinates, $r \geq 2$, each nonzero summand equals

$$\frac{\phi(1/r)^r}{\phi(1/(r-1))^{r-1}} = \frac{1/r!}{1/(r-1)!} = \frac{1}{r},$$

and vertices give equality as well. □

## 4. The permanent inequality

To deduce Theorem 1.1 from Theorem 1.3, we use Hölder's inequality to separate the rows. Each resulting factor is a sum of values of $\operatorname{per}\phi$ on normalized minors, weighted by the summands $q_{ij}$ of (1.5). Induction bounds these values, and Theorem 1.3 bounds the sum of their weights.

### 4.1. **Induction via Hölder's inequality.**

Let $M_n$ be the maximum of $\operatorname{per}\phi[A]$ over row-stochastic $n \times n$ matrices, which exists by compactness. For $n \geq 2$ and a positive row-stochastic matrix $A$, put

$$q_{ij} = \frac{\phi(a_{ij})^n}{\prod_{\ell \neq j} \phi\Big(\dfrac{a_{ij}}{1-a_{i\ell}}\Big)}.$$

Then $\sum_j q_{ij} = Q_n(a_i)$, where $a_i$ is row $i$. For $\sigma \in \mathfrak{S}_n$, define

$$W(\sigma) = \prod\nolimits_k \phi(a_{k,\sigma(k)}), \qquad V_i(\sigma) = \prod\nolimits_{k \neq i} \phi\left(\frac{a_{k,\sigma(k)}}{1-a_{k,\sigma(i)}}\right).$$

For a fixed row $k$, the values $\sigma(i)$, $i \neq k$, run over all columns except $\sigma(k)$, so

$$W(\sigma)^n = \prod\nolimits_{i=1}^n (q_{i,\sigma(i)} V_i(\sigma)).$$

Hölder's inequality gives

$$\operatorname{per}\phi[A] \leq \prod\nolimits_{i=1}^n \Big(\sum\nolimits_{\sigma \in \mathfrak{S}_n} q_{i,\sigma(i)} V_i(\sigma)\Big)^{1/n}. \tag{4.1}$$

For fixed $i, j$, the sum of $V_i(\sigma)$ over $\sigma(i) = j$ is $\operatorname{per}\phi[A^{(i,j)}]$ for the matrix

$$A^{(i,j)}_{k\ell} = \frac{a_{k\ell}}{1-a_{kj}} \qquad (k \neq i,\ \ell \neq j),$$

which is row-stochastic. Therefore

$$\operatorname{per}\phi[A] \leq M_{n-1} \prod\nolimits_{i=1}^n Q_n(a_i)^{1/n} \leq M_{n-1}.$$

Approximation by positive matrices and Section 3.4 extend the first inequality to every row-stochastic matrix. Since $M_1 = 1$, induction gives $M_n \leq 1$, proving Theorem 1.1. In particular,

$$\operatorname{per}\phi[A] \leq \prod\nolimits_{i=1}^n Q_n(a_i)^{1/n}. \tag{4.2}$$

4.2. **The equality cases.** Every block $J_r/r$ in Theorem 1.1 has $\operatorname{per}\phi[J_r/r] = r!\phi(1/r)^r = 1$, so their direct sums give equality. We prove necessity by induction on $n$; the case $n = 1$ is immediate.

Suppose $\operatorname{per}\phi[A] = 1$. By (4.2) and Theorem 1.3, every row is uniform on its positive support. Let $S_i$ be that support and $r_i = |S_i|$.

If $r_i = 1$, let $j$ be its unique column. Expansion along row $i$ shows that the minor $A'$ obtained by deleting row $i$ and column $j$ has $\operatorname{per}\phi[A'] = 1$, and its surviving row sums $1 - a_{kj}$ are positive. If some $a_{kj} > 0$, normalizing the surviving rows strictly increases that positive permanent, because every positive matching term uses a positive entry in row $k$ and $\phi$ is strictly increasing. Such an increase contradicts the inequality already proved. Thus column $j$ has no other positive entries, and an isolated $1 \times 1$ block can be removed.

It remains to consider $r_i \geq 2$ for every row. All entries are then at most $1/2$, so the Hölder argument applies directly even when some entries vanish. On row $i$,

$$q_{ij} = \begin{cases} 1/r_i, & j \in S_i, \\ 0, & j \notin S_i. \end{cases}$$

Each factor on the right of (4.1) is at most 1, so equality of the full permanent forces

$$\operatorname{per}\phi[A^{(i,j)}] = 1 \qquad (j \in S_i). \tag{4.3}$$

By induction, the support graph of every such minor is a disjoint union of complete balanced bipartite graphs, one for each block of Theorem 1.1.

Choose a row $i$ of minimum degree $r = \min_k r_k$, and choose $j \in S_i$. Let $G$ be the bipartite support graph of $A$. In $G - i - j$, every remaining row has degree at least $r - 1$, so every complete balanced block has size at least $r - 1$. Column $j$ has another neighbor $k \neq i$: otherwise an edge $i\ell$, $\ell \in S_i \setminus \{j\}$, could not extend to a perfect matching, contrary to (4.3).

In $G - k - j$, row $i$ has degree $r - 1$, so it belongs to a complete balanced block of size $r - 1$. For $\ell \in S_i \setminus \{j\}$, let the block containing $\ell$ in $G - i - j$ have size $m \geq r - 1$. Its column degree in $G - k - j$ is $m$ if $k$ belongs to that block and $m + 1$ otherwise, and it must equal $r - 1$. Thus $m = r - 1$ and $k$ belongs to the block, so all $r - 1$ columns in $S_i \setminus \{j\}$ form the same block $C$, with $r - 1$ rows. Every row of $C$ must also meet column $j$, since its original degree is at least $r$. No row outside $C \cup \{i\}$ can meet $j$: deleting such a row and $j$ would leave every column of $C$ with degree $r$, whereas the surviving row $i$ must lie in a block of size $r - 1$. Thus $C$, row $i$, and column $j$ form an isolated complete balanced block of size $r$, whose stochastic matrix is $J_r/r$. Removing it and continuing proves the equality case of Theorem 1.1.

The normalization in (1.2) converts the stochastic inequality into a bound for nonnegative rows.

***Proof of Corollary 1.2.*** The case of a zero row is immediate. Otherwise let $r_i = \|b_i\|_\psi$ and set

$$a_{ij} = \psi(b_{ij}/r_i).$$

Then $A$ is row-stochastic and $\phi(a_{ij}) = b_{ij}/r_i$. Hence

$$\frac{\operatorname{per} B}{\prod_i r_i} = \operatorname{per}\phi[A] \leq 1,$$

and the equality classification follows from Theorem 1.1. □

4.3. **A strict form.** In the relative interior, the proof of Theorem 1.3 discards the nonnegative terms $a_i a_j(2 - R_{a_j}(a_i) - R_{a_i}(a_j))$ in the last step of (3.5). Keeping them gives a strict form of Theorem 1.1 that is computable from one-dimensional integrals.

For positive $a, b$ with $a + b \leq 1$, define

$$D(a, b) = 2 - R_b(a) - R_a(b).$$

By Lemma 3.1, this symmetric function is nonnegative and vanishes exactly when $a = b$. For a probability vector $a$, put

$$\mathfrak{d}(a) = \sum_{i<j} a_i a_j D(a_i, a_j),$$

where terms with a zero coordinate are defined to be zero.

**Proposition 4.1** (Explicit strict refinement)**.** *For every $a \in \Delta_n$,*

$$Q_n(a) \le 1 - \mathfrak{d}(a). \tag{4.4}$$

*Consequently every row-stochastic $A$ satisfies*

$$\operatorname{per} \phi[A] \le \prod_{i=1}^{n} (1 - \mathfrak{d}(a_i))^{1/n} \le \exp\left\{-\frac{1}{n} \sum_{i=1}^{n} \mathfrak{d}(a_i)\right\}. \tag{4.5}$$

*If $0 < a < b$, $a + b \le 1$, and $u = (b - a)K(a, b)$, then*

$$D(a, b) \ge \frac{1}{1+u} - e^{-u} \ge \frac{u^2}{2(1+u)(1+u+u^2/2)}. \tag{4.6}$$

*Proof.* Keeping these terms in (3.5) gives (4.4); then (4.2) and $\log(1-t) \le -t$ give (4.5). For (4.6), put $v = (b-a)K(b,a)$. The proof of Lemma 3.1 gives

$$e^v \le \frac{1}{1-v} \le 1 + \frac{u}{1+u}.$$

Thus $2 - e^{-u} - e^v \ge 1/(1+u) - e^{-u}$, and $e^u \ge 1 + u + u^2/2$ gives the rational lower bound. Boundary statements follow from $0 \le D(a,b) \le 2$ and Section 3.4. □

The deficit is computed row by row. If $\operatorname{per} \phi[A] \ge 1 - \varepsilon > 0$, for example, then

$$\frac{1}{n} \sum_i \mathfrak{d}(a_i) \le -\log(1 - \varepsilon),$$

so near equality forces the average row deficit to be small.

## 5. Consequences

The formula (1.4) and the entropy functional dual to the Orlicz norm give consequences for weighted matchings and random permutations.

5.1. **Brégman–Minc, rectangular permanents, and matchings.** By (1.4), Corollary 1.2 contains the Brégman–Minc inequality with its equality cases, and its proof does not use Brégman's theorem. On a general nonnegative matrix, the bound (1.3) sits below the Brégman–Minc bound scaled by the row maxima. Suppose every row is nonzero, with support size $r_i$ and maximum $m_i$. Coordinatewise monotonicity gives

$$\operatorname{per} B \le \prod_i \|b_i\|_\psi \le \prod_i m_i (r_i!)^{1/r_i}.$$

The comparison between row factors is strict unless all positive entries in that row are equal. Thus the Orlicz norm retains information about the entries beyond their support and maximum. Transposition gives a corresponding column bound, and for a real or complex matrix the elementary inequality $|\operatorname{per} B| \le \operatorname{per} |B|$ extends the norm bound to absolute permanents. Column scaling multiplies $\operatorname{per} B$ by $\prod_j y_j$, so Corollary 1.2 also gives $\operatorname{per} B \le \inf_{y>0} \prod_i \|b_i \odot y\|_\psi / \prod_j y_j$, where $\odot$ is the coordinatewise product; a supplement to this paper shows that the logarithm of this bound is a convex program with an exact entropy dual.

Appending all-ones rows allows us to apply Corollary 1.2 to a rectangular matrix, with an explicit cost determined by $\|\mathbf{1}_n\|_\psi = (n!)^{1/n}$. For an $m \times n$ matrix $B$ with $m \le n$, its permanent is

$$\operatorname{per}_{\mathrm{inj}} B = \sum_{\sigma:[m]\hookrightarrow[n]} \prod_{i=1}^{m} b_{i,\sigma(i)}.$$

**Theorem 5.1** (Sharp rectangular bound)**.** *For every nonnegative $m \times n$ matrix $B$, $m \le n$,*

$$\operatorname{per}_{\mathrm{inj}} B \le \frac{(n!)^{(n-m)/n}}{(n-m)!} \prod_{i=1}^{m} \|b_i\|_\psi. \tag{5.1}$$

*The constant is best possible. If $m < n$ and every row is nonzero, equality holds if and only if each row is a positive constant across all $n$ columns.*

*Proof.* Append $n-m$ all-ones rows to obtain a square matrix $\widetilde{B}$. Every injection $[m] \hookrightarrow [n]$ extends to exactly $(n-m)!$ permutations of $[n]$, so

$$\operatorname{per} \widetilde{B} = (n-m)! \operatorname{per}_{\mathrm{inj}} B.$$

Since $\|\mathbf{1}_n\|_\psi = (n!)^{1/n}$, Corollary 1.2 applied to $\widetilde{B}$ proves (5.1), and constant rows give equality. If $m < n$, an appended all-ones row has full support, so the equality cases of Corollary 1.2 force a single block of size $n$, proving the equality assertion. □

For an arbitrary nonnegative $m \times n$ matrix and $0 \le k \le \min(m,n)$, let

$$M_k(B) = \sum_{\substack{I \subseteq [m],\ J \subseteq [n] \\ |I|=|J|=k}} \operatorname{per} B[I,J], \qquad M_0(B) = 1,$$

the total weight of all $k$-matchings in the bipartite graph with edge weights $b_{ij}$. Bounding every $M_k(B)$ also bounds the weighted rook polynomial $\sum_k M_k(B) z^k$ at every $z \ge 0$. Let $e_k$ denote the $k$-th elementary symmetric polynomial.

**Corollary 5.2** (Coefficientwise matching bound)**.** *Put $r_i = \|b_i\|_\psi$. Then*

$$M_k(B) \le \frac{(n!)^{(n-k)/n}}{(n-k)!} e_k(r_1, \dots, r_m) \qquad (0 \le k \le \min(m,n)).$$

*All these coefficient bounds are attained simultaneously by matrices whose rows are constant.*

*Proof.* For a fixed $k$-element row set $I$, the sum over the $k$-element column sets $J$ is $\operatorname{per}_{\mathrm{inj}} B[I,:]$. Apply Theorem 5.1 and sum over $I$. If $b_{ij} = c_i$, simultaneous equality follows from

$$M_k(B) = \frac{n!}{(n-k)!} e_k(c_1, \dots, c_m), \qquad r_i = (n!)^{1/n} c_i.$$ □

5.2. **Entropy of random permutations.** The entropy bound from the introduction follows by optimizing the Gibbs variational inequality over the rows of a stochastic matrix. This optimization defines the entropy functional dual to the Orlicz norm. For a probability vector $p \in \Delta_n$, define

$$\mathcal{H}_\Gamma(p) = \inf_{a \in \Delta_n} \sum\nolimits_j p_j h(a_j), \qquad E_\Gamma(p) = \sum\nolimits_j p_j h(p_j) = \sum\nolimits_j p_j^2 \log \Gamma(1 + 1/p_j).$$

All logarithms are natural. The convention is $0 \cdot h(0) = 0$, and a positive coefficient multiplying $h(0)$ gives $+\infty$. Zero coordinates can equivalently be deleted, since $h$ is decreasing and putting positive mass on a zero coordinate of $p$ cannot improve the infimum.

**Proposition 5.3** (Properties of $\mathcal{H}_\Gamma$ and $E_\Gamma$)**.** *The function $\mathcal{H}_\Gamma$ is symmetric and concave. The function $E_\Gamma$ is symmetric and strictly concave on each nontrivial simplex. Moreover,*

$$0 \le \mathcal{H}_\Gamma(p) \le E_\Gamma(p) \le \mathrm{H}(p), \qquad \mathrm{H}(p) = -\sum\nolimits_j p_j \log p_j.$$

*If $p$ has support size $r$, then*

$$E_\Gamma(p) \le \frac{\log(r!)}{r}, \tag{5.2}$$

*strictly unless $p$ is uniform on its support. At that uniform vector, both $\mathcal{H}_\Gamma$ and $E_\Gamma$ equal $\log(r!)/r$.*

*Proof.* Approximating boundary competitors $a$ by positive vectors shows that the infimum defining $\mathcal{H}_\Gamma$ can be restricted to the relative interior of the simplex. It is therefore an infimum of finite affine functions of $p$, which proves concavity. By (2.1), $xh(x) = -\gamma x + \int_0^1 \omega(t)\, x/(x+t)\, dt$, and hence

$$\frac{d^2}{dx^2}(xh(x)) = -2\int_0^1 \frac{\omega(t)t}{(x+t)^3}\, dt < 0.$$

The function $xh(x)$ extends continuously to zero with value zero. Hence $E_\Gamma$ is strictly concave, and Jensen's inequality yields (5.2). Testing $a = p$ proves $\mathcal{H}_\Gamma \le E_\Gamma$. Since $\phi(x) \ge x$ on $[0,1]$, $h(x) \le -\log x$, giving $E_\Gamma \le \mathrm{H}$, and nonnegativity follows from $h \ge 0$ on $(0,1]$. For a uniform vector, convexity of $h$ proves that the infimum defining $\mathcal{H}_\Gamma$ occurs at the same uniform vector. □

**Theorem 5.4** (Entropy of a random permutation)**.** *For a random permutation $\Sigma$ of $[n]$, put*

$$p_{ij} = \mathbb{P}\{\Sigma(i) = j\}, \qquad p_i = (p_{ij})_j.$$

*Then, with $\mathrm{H}(\Sigma)$ the Shannon entropy of the law of $\Sigma$,*

$$\mathrm{H}(\Sigma) \le \sum\nolimits_i \mathcal{H}_\Gamma(p_i) \le \sum\nolimits_i E_\Gamma(p_i) = \sum\nolimits_{i,j} p_{ij}^2 \log\Gamma(1 + 1/p_{ij}). \tag{5.3}$$

*The first inequality is equivalent to Theorem 1.1. The bounds are sharp for independent uniform permutations within disjoint blocks.*

*Proof.* For a positive row-stochastic $A$, let

$$w(\sigma) = \prod\nolimits_i \phi(a_{i,\sigma(i)}), \qquad Z = \sum\nolimits_\sigma w(\sigma) \le 1.$$

Nonnegativity of the relative entropy of the law of $\Sigma$ with respect to $w/Z$ gives

$$\mathrm{H}(\Sigma) \le \sum\nolimits_{i,j} p_{ij} h(a_{ij}) + \log Z \le \sum\nolimits_{i,j} p_{ij} h(a_{ij}).$$

Taking the infimum over each row independently, with approximation to permit zero coordinates in the minimizers, gives the first bound, and Proposition 5.3 gives the second.

Conversely, if the first bound holds, apply it to the distribution $w/Z$. Its entropy identity yields

$$\log Z = \mathrm{H}(w/Z) - \sum\nolimits_{i,j} p_{ij} h(a_{ij}) \le \sum\nolimits_i \left( \mathcal{H}_\Gamma(p_i) - \sum\nolimits_j p_{ij} h(a_{ij}) \right) \le 0.$$

Since $Z = \operatorname{per}\phi[A]$, this bound on $\log Z$ is the permanent bound for positive $A$, and continuity handles the boundary. For a block of size $r$, a uniform permutation has entropy $\log(r!)$, and its $r$ row contributions are each $\log(r!)/r$; independent blocks therefore give equality. □

Counting is the special case of a uniform law on a family of permutations, where the entropy is the logarithm of the family's size.

**Corollary 5.5** (A bound using marginal frequencies)**.** *Let $\mathcal{F}$ be a nonempty family of permutations, and let $p_{ij}$ be the marginal frequencies of the uniform distribution on $\mathcal{F}$. Then*

$$|\mathcal{F}| \le \prod\nolimits_{i,j:\, p_{ij}>0} \Gamma(1 + 1/p_{ij})^{p_{ij}^2}.$$

*If row $i$ uses $r_i$ possible columns, this bound is at most $\prod_i (r_i!)^{1/r_i}$, and is strictly smaller whenever some row marginal is nonuniform on its support.*

*Proof.* Use $\mathrm{H}(\Sigma) = \log|\mathcal{F}|$ in Theorem 5.4, followed by (5.2). □

Entropy methods for Brégman's inequality go back to Radhakrishnan (1997); Corollary 5.5 uses the marginal frequencies to refine the bound based on support sizes.

## AI USAGE

The results were developed with GPT-6 (Pro) after extensive back and forth; none were obtained zero-shot. The author chose the problem, directed its Orlicz formulation, and suggested seeking positive integral representations as in subentropy (Sra, 2019), which was decisive to unlocking the proof—Section 3. GPT Pro supplied derivations, literature searches, and repeated checks; GPT and Claude helped polish the writing.

## ACKNOWLEDGEMENTS

The author (SS) thanks the Alexander von Humboldt (AvH) Foundation for their generous support via a Humboldt Professorship in AI.

## REFERENCES

L. M. Brégman. Some properties of nonnegative matrices and their permanents. *Soviet Math. Dokl.*, 14:945–949, 1973. Translated from Dokl. Akad. Nauk SSSR **211** (1973), no. 1, 27–30.

Alexandre Bristiel and Pietro Caputo. Entropy inequalities for random walks and permutations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 60(1):54–81, 2024. doi: 10.1214/22-AIHP1267. Preprint: arXiv:2109.06009.

Leonid Gurvits and Alex Samorodnitsky. Bounds on the permanent and some applications. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 90–99. IEEE Computer Society, 2014. doi: 10.1109/FOCS.2014.18. Extended version: arXiv:1408.0976; conjecture numbering refers to that version.

Richard Jozsa and Graeme Mitchison. Symmetric polynomials in information theory: entropy and subentropy. *J. Math. Phys.*, 56(6):062201, 2015. doi: 10.1063/1.4922317. Preprint: arXiv:1404.7694.

National Institute of Standards and Technology. NIST digital library of mathematical functions. `https://dlmf.nist.gov/5.8`, 2026. Chapter 5, especially §5.8. Accessed 30 October 2026.

Jaikumar Radhakrishnan. An entropy proof of Brégman's theorem. *J. Combin. Theory Ser. A*, 77 (1):161–164, 1997. doi: 10.1006/jcta.1996.2727.

Alex Samorodnitsky. An upper bound for permanents of nonnegative matrices. *J. Combin. Theory Ser. A*, 115(2):279–292, 2008. doi: 10.1016/j.jcta.2007.05.010. Preprint: arXiv:math/0605147.

George W. Soules. Permanental bounds for nonnegative matrices via decomposition. *Linear Algebra Appl.*, 394:73–89, 2005. doi: 10.1016/j.laa.2004.06.022.

Suvrit Sra. Logarithmic inequalities under a symmetric polynomial dominance order. *Proc. Amer. Math. Soc.*, 147(2):481–486, 2019. doi: 10.1090/proc/14023. Preprint, with a slightly different title: arXiv:1509.05902.