An explicit Frobenius corollary of established marginal entropy bounds for permanents

Alec Kriebel
ORCID: 0009-0001-9320-500X

Version 1.0, 14 October 2026

Abstract

We give a self-contained, credited corollary of the Gibbs-marginal random-order comparison of Anari and Rezaei. For every nonnegative row-stochastic matrix A of order $n \geq 1$, writing $S = \sum_{i,j} a_{ij}^2$, the permanent is at most $\exp\{-n + e^4 S(1 + \log n)\}$. The proof combines conditional Jensen with an explicit scalar relative-entropy estimate. For fixed $S \leq \gamma$, it supplies a polynomial-factor e^{-n} upper bound for the class in the Barvinok–Samorodnitsky question recorded in OWR 44/2008. The same corollary follows from Anari's 2026 sequential identity and Dong–Jain's matching entropy theorem. The estimate is not sharp; historical novelty and priority of the scalar or final norm statement remain unknown.

1 Statement, source question, and credit

All logarithms are natural. Put $h(a) = a^2 \log(e/a)$ for $a > 0$ and $h(0) = 0$. A row-stochastic matrix is nonnegative with every row sum equal to one; it need not have column sums equal to one.

Corollary 1. *For every row-stochastic $n \times n$ matrix A, $n \geq 1$, and $S = \sum_{i,j} a_{ij}^2$,*

$$\operatorname{per}(A) \leq \exp\left\{-n + e^4 \sum_{i,j} h(a_{ij})\right\} \leq \exp\{-n + e^4 S(1 + \log n)\}. \tag{1}$$

If $S \leq \gamma$, then for all $n \geq 1$,

$$\operatorname{per}(A) \leq e^{e^4 \gamma} n^{e^4 \gamma} e^{-n}. \tag{2}$$

For $n \geq 2$ this also gives $\operatorname{per}(A) \leq n^{3e^4\gamma} e^{-n}$.

Problem 9 of Barvinok and Samorodnitsky [1, pp. 2549–2550] fixes $\gamma \geq 1$ and bounds the *unnormalized* squared Frobenius norm $S \leq \gamma$ for doubly stochastic matrices. Its polynomial-factor exponential scale follows from (2) and the already-known van der Waerden lower bound $n!/n^n$ recorded there. We claim no lower bound for general row-stochastic inputs. The source's stronger hypothesis $a_{ij} \leq \gamma/n$ is not substituted for $S \leq \gamma$. At $n = 1$, $\operatorname{per}([1]) = 1$, so a prefactor-free $n^C e^{-n}$ upper bound is false.

The direct precursor is Anari–Rezaei [2, §4, Eq. (6), p. 12], using Gibbs marginals rather than input entries and random exposure orders. The classical entropy method is credited to Radhakrishnan [3]. Sections 2–3 reconstruct that mechanism and its Jensen simplification, including sparse support. Section 4 supplies an explicit scalar estimate and the norm conversion. They are estimates written out here; no new-method, first-solution, or independent-discovery claim is made. Section 5 explains further affirmative source implications.

2 Random exposure and conditional Jensen

Define

$$q(b) = \frac{-b^2 \log b}{1-b} \quad (0 < b < 1), \qquad q(0) = 0, \quad q(1) = 1.$$

The endpoint values are the continuous limits.

Lemma 2. *If $X = (X_1, \dots, X_n)$ is any random permutation and $b_{ij} = \Pr(X_i = j)$, then*

$$H(X) \le -\sum_{i,j} b_{ij} \log b_{ij} - n + \sum_{i,j} q(b_{ij}). \tag{3}$$

Here $0 \log 0 = 0$.

Proof. Independently of X, choose independent uniform priorities $T_i \in [0,1]$ and expose rows in increasing order. Ties have probability zero. Fix an order and a preceding history of positive probability. Let U_i be the unused columns, c_j the conditional law of X_i, and $Z_i = \sum_{j \in U_i} b_{ij}$. Whenever $c_j > 0$, $j \in U_i$ and $b_{ij} > 0$; hence $Z_i > 0$. The law $r_j = b_{ij}/Z_i$ on U_i contains the support of c. Thus

$$\mathrm{KL}(c \| r) = \sum_j c_j \log(c_j / b_{ij}) + \log Z_i \ge 0.$$

Averaging $H(c) + \sum_j c_j \log b_{ij} \le \log Z_i$ over positive histories and independent orders, the entropy chain rule gives

$$H(X) + \sum_{i,j} b_{ij} \log b_{ij} \le \sum_i \mathbb{E} \log Z_i. \tag{4}$$

No conditional law on an impossible history is used.

Condition on a supported full permutation X and $T_i = t$. Since X is bijective, the available columns are X_i and those X_k with $T_k > t$. Independence and the marginal row sum give

$$\mathbb{E}[Z_i \mid X, T_i = t] = b_{iX_i} + (1-t)(1 - b_{iX_i}).$$

For every sampled permutation $b_{iX_i} > 0$, and $Z_i \ge b_{iX_i}$. Finite positive support bounds the sampled logarithms, justifying these averages. Concavity of log and, for $0 < b < 1$, the integral

$$\int_0^1 \log\{b + (1-t)(1-b)\}\, dt = -1 - \frac{b \log b}{1-b}$$

yield $\mathbb{E} \log Z_i \le -1 + \sum_j q(b_{ij})$. The integral is zero at $b = 1$; zero marginal values are never sampled. The second factor b in q comes from averaging over X_i. Substitute in (4). A deterministic permutation gives equality and checks $q(1) = 1$. □

3 Gibbs specialization and matrix relative entropy

If $\mathrm{per}(A) = 0$, (1) holds immediately. Otherwise define the law

$$\mu(\sigma) = \frac{\prod_i a_{i\sigma(i)}}{\mathrm{per}(A)}$$

on permutations, and let $B = (b_{ij})$ be its marginals. Then B is doubly stochastic and $b_{ij} > 0$ implies $a_{ij} > 0$. In general $B \neq A$. The exact entropy identity and Lemma 2 give

$$\begin{aligned} \log \operatorname{per}(A) &= H(\mu) + \sum_{i,j} b_{ij} \log a_{ij}, \\ \log \operatorname{per}(A) &\leq -n + \sum_{i,j} q(b_{ij}) - D_{\mathrm{mat}}(B\|A), \\ D_{\mathrm{mat}}(B\|A) &= \sum_{b_{ij}>0} b_{ij} \log(b_{ij}/a_{ij}). \end{aligned} \tag{5}$$

This is the matrix relative entropy: $n\,\mathrm{KL}(B/n\|A/n)$. It is distinct from the sequential permutation divergence discussed in Section 5.

Both matrices have total mass n, so

$$D_{\mathrm{mat}}(B\|A) = \sum_{i,j} d(b_{ij}, a_{ij}), \qquad d(b,a) = b\log(b/a) - b + a. \tag{6}$$

For $a > 0$, $d(b,a) \geq 0$ by $x \log x - x + 1 \geq 0$. Its boundary values are $d(0,a) = a$ and $d(0,0) = 0$; $a = 0 < b$ cannot occur. Input-positive, marginal-zero entries must retain their a term in (6). Only the *total* linear terms cancel. Column stochasticity of A is unnecessary.

4 Explicit scalar estimate and Frobenius conversion

Lemma 3. *For $0 \leq a, b \leq 1$, with $a = 0$ permitted only at $b = 0$,*

$$q(b) - d(b,a) \leq e^4 h(a). \tag{7}$$

Proof. At $(a,b) = (0,0)$ the claim is immediate; at $b = 0 < a$ the left side is $-a$. For $0 < b < 1$, $b\log(1/b) \leq 1 - b$ gives both $q(b) \leq b$ and $q(b) \leq h(b)$. The endpoint comparisons follow by continuity. Also

$$h'(b) = b\{1 + 2\log(1/b)\} > 0 \qquad (0 < b \leq 1).$$

If $a, b > 0$ and $b \geq e^2 a$, then $\log(b/a) \geq 2$, so $d(b,a) \geq b + a \geq q(b)$. If $b < e^2 a$, discard $d \geq 0$. When $b \leq a$, monotonicity on $[0,1]$ gives $h(b) \leq h(a)$. When $a < b < e^2 a$, use $b^2 < e^4 a^2$ and $\log(e/b) \leq \log(e/a)$ separately. Thus $h(b) \leq e^4 h(a)$. This includes $b = 1$ by its endpoint value and the switch line in the first case. Monotonicity is used only between points in $[0,1]$. □

Summing (7) in (5)–(6) proves the first bound of (1). Row Cauchy–Schwarz gives $S \geq 1$. The weights $p_{ij} = a_{ij}^2/S$ form a probability law on at most n^2 points, so $H(p) \leq 2\log n$ and

$$\sum_{i,j} a_{ij}^2 \log(1/a_{ij}) = \frac{S}{2}\{H(p) - \log S\} \leq S \log n.$$

Zero terms use their continuous values. Adding S proves the second bound. For $n \geq 2$, $1 + \log n \leq 3\log n$, proving the prefactor-free consequence.

Some polynomial loss is necessary. For a fixed integer $k \geq 1$ with $k \leq \gamma$, take $n = km$ and k uniform blocks J_m/m. Then $S = k$ and Stirling's formula gives

$$\operatorname{per}(A) = (m!/m^m)^k \sim (2\pi/k)^{k/2} n^{k/2} e^{-n}.$$

An exponent valid at $\gamma = k$ is therefore at least $k/2$ asymptotically. The exponent $3e^4\gamma$ is not claimed sharp. Fixed γ is essential to a fixed-degree polynomial conclusion; $S/n \leq \gamma$ would instead admit permutation matrices.

5 Further source implications and limits

For a probability vector b, let

$$T(b) = \sum_j b_j \mathbb{E}_\pi \log \left(\sum_{k \text{ at or after } j \text{ in } \pi} b_k \right),$$

where π is a uniform coordinate order. Anari–Rezaei's cited comparison is

$$\log \operatorname{per}(A) \le -D_{\mathrm{mat}}(B\|A) + \sum_i T(B_i).$$

Conditional Jensen, as above, yields $T(b) \le -1 + \sum_j q(b_j)$, giving (5). For source statements on positive inputs, one can also approximate by $(1-\varepsilon)A+\varepsilon J_n/n$ and pass to the continuous input-only bound.

Anari [4, §4, Lemma 7, Eq. (24), pp. 9–10] retains the sequential term exactly. Expanding that identity on positive inputs gives

$$\log \operatorname{per}(A) = -D_{\mathrm{mat}}(B\|A) + \sum_i T(B_i) - D_{\mathrm{seq}},$$
$$D_{\mathrm{seq}} = \mathbb{E}_\pi \operatorname{KL}(\mu\|\nu_\pi) \ge 0.$$

Here ν_π is the sequential marginal reference law on permutations. D_{seq} compares *permutation distributions*, whereas D_{mat} compares *normalized entry masses*. They cannot be identified. The source's vector function named g is also distinct from our scalar q. Discarding D_{seq}, applying Jensen, and then Lemma 3 proves the same corollary.

Dong–Jain [5, §3, Theorem 3.1, p. 9, $g = 4$] supplies another route to Lemma 2. On the complete bipartite graph, put

$$H_{\mathrm{B}}(B) = -\sum b_{ij} \log b_{ij} + \sum (1 - b_{ij}) \log(1 - b_{ij}),$$
$$F(b) = T(b) - \sum_j (1 - b_j) \log(1 - b_j).$$

Their specialization bounds $H(\mu)$ by $H_{\mathrm{B}}(B) + \frac{1}{2}\sum_v F(B_v)$. Each edge occurs at both endpoints; expansion cancels the $(1 - b)\log(1 - b)$ terms. The result is

$$H(\mu) \le -\sum b_{ij} \log b_{ij} + \tfrac{1}{2} \left\{ \sum_i T(B_i) + \sum_j T(B^j) \right\}.$$

Jensen on rows and columns gives exactly (3). These affirmative implications establish credited dependence, without deciding whether the final scalar or norm formula was explicitly published earlier.

A bounded source audit dated 14 October 2026 did not settle historical priority. Full Soules bodies from 2000, 2003, and 2005 were inaccessible (HTTP 403); richer Gurvits–Samorodnitsky Orlicz, Roos subpermanent, and optimized ℓ_p consequences were not exhaustively classified. Their broader implications remain unknown. No exhaustive literature, worldwide open-status, or novelty certificate is asserted.

Verification and preparation record. The proof above is analytic. The companion kit contains exact copies of nine historical submission files, distinct full mathematical/source reviews, and finite diagnostic programs. Numerical checks corroborate selected cases; they are not interval certificates or universal proofs. This version uses extensive AI assistance, including drafting, computations, and adversarial analysis. It is unrefereed, with no external human peer review or proof-assistant certification. The dated preparation snapshot is in the companion kit's `STATUS.json`; the separate delivery record `FINAL_VERIFICATION.json` records integrity checks. Publication decisions are separate from mathematical statements. Copyright Alec Kriebel, 2026, licensed CC BY 4.0.

References

[1] A. Barvinok and A. Samorodnitsky, Problem 9, in *Discrete Geometry*, Oberwolfach Reports 5 (2008), no. 4, 2549–2550. Report 44/2008; publisher publication date 30 September 2009. doi:10.4171/OWR/2008/44; official PDF file 46191.

[2] N. Anari and A. Rezaei, *A Tight Analysis of Bethe Approximation for Permanent*, arXiv:1811.02933v2, revised 9 December 2019. Formula comparison uses §4, Eq. (6), p. 12.

[3] J. Radhakrishnan, *An Entropy Proof of Bregman's Theorem*, J. Combin. Theory Ser. A 77 (1997), 161–164. doi:10.1006/jcta.1996.2727. Classical attribution; no uninspected theorem is needed above.

[4] N. Anari, *Beyond the Bethe Approximation of the Permanent*, arXiv:2608.28031v2, revised 1 September 2026. §4, Lemma 7, pp. 9–10.

[5] D. Dong and V. Jain, *Optimal Girth-Dependent Bounds for the Bethe Approximation of the Permanent*, arXiv:2609.02017v1, 2 September 2026. §3, Theorem 3.1, p. 9.