The sharp top-mean index in dimensions at least five

Research proof and formalization packet
October 26, 2026

Abstract

For a probability vector in dimension d, form all d^2 arithmetic means and all d^2 geometric means of its ordered coordinate pairs. Sort the two arrays independently. We prove that, for every $d \geq 5$, the sum of their respective largest k entries is at most d for every probability vector if and only if $k \leq 2d-5$. This establishes the value proposed by Caro and Graswald in that dimension range. The proof combines two incidence bounds for arbitrary sets of ordered pairs with Cauchy–Schwarz. An explicit probability vector proves sharpness in every dimension. The complete statement, including the independently sorted arrays, is certified in Lean.

mathdb:2610.00775v1 28 Oct 2026

1 Statement and source

Let $p = (p_1, \dots, p_d)$ satisfy $p_i \geq 0$ and $\sum_i p_i = 1$. Set

$$a_{ij} = \frac{p_i + p_j}{2}, \qquad g_{ij} = \sqrt{p_i p_j} \qquad (1 \leq i, j \leq d).$$

Write $a_1^{\downarrow}, \dots, a_{d^2}^{\downarrow}$ and $g_1^{\downarrow}, \dots, g_{d^2}^{\downarrow}$ for these two arrays in nonincreasing order, retaining repetitions, and put

$$A_k(p) = \sum_{\ell=1}^{k} a_\ell^{\downarrow}, \qquad G_k(p) = \sum_{\ell=1}^{k} g_\ell^{\downarrow}.$$

The selected ordered pairs may differ between the two sums. Caro–Graswald [1], Conjecture 4.19 in arXiv v3 (Conjecture IV.19 in the published paper), propose the maximal index $h(d) = 2d-5$ for the universal inequality $A_{h(d)} + G_{h(d)} \leq d$. Their Remark 4.18 explains its role in necessary criteria for infinitesimal Markovian divisibility. We prove the proposed value for every $d \geq 5$.

Theorem 1. *Let $d \geq 5$ and $0 \leq k \leq d^2$ be integers. Then*

$$\big[A_k(p) + G_k(p) \leq d \text{ for every probability vector } p\big] \quad \Longleftrightarrow \quad k \leq 2d-5.$$

Moreover, $A_{2d-5}(p) + G_{2d-5}(p) < d$ for every such p.

For a nonnegative finite array, its largest-k sum is the maximum of the sums over k distinct indices. Thus upper bounds may be proved for arbitrary sets of ordered pairs, and lower bounds by exhibiting such sets. This also handles ties and zero entries.

2 The universal bound

Put $m = d-3$ and $k = 2m+1$. Write $P = \max_i p_i$. For any set S of k ordered pairs define

$$c_i = \frac{\#\{j : (i,j) \in S\} + \#\{j : (j,i) \in S\}}{2}.$$

Then the corresponding arithmetic sum is $\sum_i c_i p_i$, and

$$c_i \leq m+1, \qquad c_i + c_j \leq m + \frac{5}{2} \quad (i \neq j). \tag{1}$$

Indeed, for one vertex the loop contributes 1 and each other incident ordered pair contributes $1/2$. For two distinct vertices, at most four selected ordered pairs have both endpoints among them; these contribute 1 to $c_i + c_j$, and every other pair contributes at most $1/2$. No symmetry assumption on S is used.

Choose u maximizing c_i, and let $a = c_u$ and $b = \max_{i \neq u} c_i$. Since $a \geq b$, the arithmetic sum is at most

$$ap_u + b(1 - p_u) \leq aP + b(1 - P).$$

If $P \leq 2/3$, this is at most

$$\frac{2a + b}{3} = \frac{a + (a + b)}{3} \leq \frac{4m + 7}{6}. \tag{2}$$

Also $\sum_{i,j} g_{ij}^2 = (\sum_i p_i)^2 = 1$, so Cauchy–Schwarz gives $G_k \leq \sqrt{k}$. Consequently

$$A_k + G_k \leq \frac{4m + 7}{6} + \sqrt{2m + 1} < m + 3.$$

For the strict inequality, the positive margin $(2m + 11)/6$ exceeds $\sqrt{2m + 1}$ because

$$(2m + 11)^2 - 36(2m + 1) = (2m - 7)^2 + 36 > 0.$$

Now suppose $P > 2/3$, and choose u with $p_u = P$. Put $\varepsilon = 1 - P < P/2$. Every arithmetic entry outside row and column u is at most ε. Give each selected entry the baseline $P/2$. The loop (u, u) needs an extra $P/2$; an entry (u, i) or (i, u) with $i \neq u$ needs an extra $p_i/2$. The loop appears at most once, and each tail coordinate appears in at most two such entries. Hence

$$A_k \leq \frac{kP}{2} + \frac{P}{2} + \varepsilon = 1 + mP. \tag{3}$$

The largest geometric entry is $g_{uu} = P$. A maximizing geometric selection can include it: if necessary replace any selected index by (u, u), which does not decrease the sum. Applying Cauchy–Schwarz to the remaining $k - 1 = 2m$ entries gives

$$G_k \leq P + \sqrt{2m}\sqrt{1 - P^2}. \tag{4}$$

Combining (3) and (4), and applying two-dimensional Cauchy–Schwarz, yields

$$\begin{aligned} A_k + G_k &\leq 1 + (m + 1)P + \sqrt{2m}\sqrt{1 - P^2} \\ &\leq 1 + \sqrt{(m + 1)^2 + 2m} = 1 + \sqrt{m^2 + 4m + 1} < m + 3. \end{aligned}$$

The final strict inequality follows from $(m + 2)^2 - (m^2 + 4m + 1) = 3$. This proves the universal assertion at $k = 2d - 5$. The same proof works for $d \geq 3$.

3 A uniform sharpness family

Let $m = d - 3 \geq 2$, $r = 2m$, and $D = r^2 + m$. Take

$$p = \frac{1}{D}\big(r^2, \underbrace{1, \ldots, 1}_{m \text{ coordinates}}, 0, 0\big).$$

At the next index $k' = 2m+2$, choose the following arithmetic entries: the distinguished diagonal, all $2m$ orientations joining the distinguished coordinate to the positive tail, and one orientation joining it to a zero coordinate. Choose the geometric entries by replacing this last arithmetic index with one positive-tail diagonal. Each selection has exactly $2m + 2$ distinct ordered pairs. Their sums are

$$A_{\text{sel}} = \frac{(m + \frac{3}{2})r^2 + m}{D}, \qquad G_{\text{sel}} = \frac{r^2 + 2mr + 1}{D}.$$

Since the independently sorted sums dominate these respective selections,

$$\begin{aligned} A_{k'} + G_{k'} - d &\geq \frac{-r^2/2 + 2mr - m^2 - 2m + 1}{r^2 + m} \\ &= \frac{(m - 1)^2}{4m^2 + m} > 0. \end{aligned}$$

For example, $d = 5$ gives $p = (16, 1, 1, 0, 0)/18$ and selected sums $29/9$ and $11/6$, whose total is $91/18 > 5$. All entries are nonnegative, so both largest-k sums are nondecreasing in k. The witness excludes every $k \geq 2d - 4$; the universal bound proves every $k \leq 2d - 5$. This completes the proof of Theorem 1.

mathdb:2610.00775v1 28 Oct 2026

4 Formal certificate and scope

The certificate comprises 76 declarations in twelve Lean modules, checked with Lean 4.34.1 and a pinned mathlib revision. Its endpoint `MaxIndex.source_index_classification` is precisely the equivalence in Theorem 1, quantified over all nonnegative functions $p : \mathrm{Fin}\, d \to \mathbb{R}$ of sum 1. The strict sufficiency theorem also covers $d \geq 3$. The development proves the literal sorted-array/finite-selection equivalence, the loop and pair incidence counts, the geometric square-mass identity, the largest-entry exchange, and the sizes and sums of both sharpness selections. Reindexing the explicit witness onto $\mathrm{Fin}\, d$ is proved as well. No combinatorial or numerical certificate remains as an extra hypothesis.

The strict replay checks every declaration type and allows only `propext`, `Classical.choice`, and `Quot.sound`. The elementary incidence and Cauchy–Schwarz ingredients are standard; the source-question conclusion is the combined sharp index theorem for $d \geq 5$. Small-dimensional classifications are outside this certificate's maximality statement. Source and bounded literature-search records accompany the development; they do not certify absolute worldwide priority.

References

[1] M. C. Caro and B. R. Graswald, *Necessary criteria for Markovian divisibility of linear maps*, J. Math. Phys. **62**, 042203 (2021), doi:10.1063/5.0031760; arXiv:2009.06666v3 (2022), Conjecture 4.19.

mathdb:2610.00775v1 28 Oct 2026