

Sharp inertia ceilings and optimal stability for inverse self-commutators

Lluís Eriksson

Independent researcher | 5 September 2026

Consolidated review revision 3

Abstract. We determine the sharp worst inverse self-commutator cost for every prescribed inertia, including arbitrary ambient zeros, and quantify its extremal geometry. For unequal sign multiplicities we obtain both optimal constants relating the normalized cost deficit to distance from the one-spike, opposite-flat boundary. The reverse coefficient is also optimal for every balanced inertia. For three positive and three negative eigenvalues we close the other direction as well: the optimal forward coefficient is $17/36$. Its proof reduces a piecewise-affine distance to 22 rational vertices and supplies exact Horn witnesses; a boundary spectrum gives ambient-independent sharpness. Thus every nearly extremal sequence is classified. A transportation refinement gives computable primal-dual upper certificates, while an exact example shows a 12.5% gap from the actual matrix cost. A sharp deficit threshold also reduces rank-ceiling stability to the known one-spike case. Classical Horn sufficiency is imported; all finite certificates are supplied for replay. The optimal balanced forward coefficient for multiplicity at least four and the general interior cost remain open here.

1. Main result and normalization

Let $F = F^* \in M_d(\mathbb{C})$ be nonzero and traceless. Use the unnormalized Hilbert--Schmidt norm and define

$$\kappa_d(F) = \frac{1}{2} \min \{ \|C\|_{HS}^2 : CC^* - C^*C = 2F \}, \quad P = \operatorname{tr} F_+ = \frac{1}{2} \|F\|_1. \quad (1)$$

Write $m = n_+(F)$, $n = n_-(F)$, $M = \max(m, n)$, and $\rho = m + n$. **Theorem 1 (sharp inertia ceiling).**

$$P \leq \kappa_d(F) \leq \frac{M+1}{2} P. \quad (2)$$

The constant $(M+1)/2$ is the exact supremum of $\kappa_d(F)/P$ on every fixed-inertia stratum $(m, n, d - m - n)$. If $m, n \geq 2$, the upper inequality is strict at every point of that stratum and its supremum is approached at a one-spike boundary. If one sign multiplicity is one, equality holds exactly when the opposite-sign eigenvalues are equal. All conclusions allow any number of ambient zero eigenvalues.

This improves the rank-only bound $\rho P/2$ in ARR-2026-1D2QV1RP1292JREW by $(\min(m, n) - 1)P/2$. The lower bound and the one-spike equality characterization are antecedents [1,2].

The minimum exists. For any traceless real diagonal target, cyclically start its eigenvalue list after a minimum partial sum. All resulting partial sums are nonnegative; a finite weighted shift with twice these partial sums as squared weights realizes the target. The

feasible set is closed, and bounded objective sublevels are compact. Unitary conjugation handles a general Hermitian target. For the zero target the cost is zero; normalized statements below concern only nonzero targets. Sign reversal preserves the cost by replacing a factor with its adjoint.

This normalization equals the predecessor's minimum of $\|H\|_{HS}\|K\|_{HS}$ over Hermitian pairs with $-i[H, K] = F$: set $C = H + iK$ and balance the two norms by the reciprocal rescaling $(H, K) \mapsto (tH, K/t)$. The identity $\|C\|_{HS}^2 = \|H\|_{HS}^2 + \|K\|_{HS}^2$ and the arithmetic-geometric mean inequality prove equivalence.

Sections 5 and 6 give optimal stability constants, the balanced three-level completion, and a certified transport refinement. This manuscript consolidates the earlier private ceiling revision and the private note on optimal constants; it is intended as one article. The exact one-spike and low-dimensional cost formulas retain their prior attribution [1,2,9,10].

2. Spectral convexity and the one-spike formula

We use the finite-dimensional Horn-Klyachko theorem [4,5]: the ordered spectra of Hermitian $A, B, A + B$ form a convex polyhedral cone. The sufficiency direction is needed, not merely necessity of individual inequalities.

Consequently, on the trace-zero ordered Weyl chamber,

$$\phi_d(\lambda) := \kappa_d(\text{diag } \lambda) \quad (3)$$

is convex and positively homogeneous. Indeed, set $s = \lambda(CC^*)/2 = \lambda(C^*C)/2$, ordered and nonnegative. Feasibility is exactly Horn feasibility of

$$(s, -s^{\text{rev}}, \lambda), \quad (4)$$

and the objective is $\sum_i s_i$. If this triple is feasible, Horn supplies isospectral positive R, S with $R - S = \text{diag } \lambda$. Choose a unitary V with $R = VSV^*$ and set $C = \sqrt{2}VS^{1/2}$. Then $CC^* = 2R$ and $C^*C = 2S$. The epigraph in (λ, s, t) , $\sum s_i \leq t$, is convex; so is its projection. The epigraph is in fact polyhedral: it is the projection of the finite Horn polyhedron with the linear trace bound. Since the cost is finite on the entire ordered trace-zero chamber, this epigraph represents the cost there as the maximum of finitely many linear forms. Hence the cost is continuous, including at the boundary of that chamber. We will use this fixed-dimensional continuity for limits that change inertia. No assertion of convexity on the whole space of Hermitian matrices is used. Such an assertion is false: the two paired targets $\text{diag}(1, -1, 0)$ and $\text{diag}(1, 0, -1)$ have cost 1 each, whereas their sum $\text{diag}(2, -1, -1)$ has cost 3.

Recalled one-spike lemma. For a target with nonzero spectrum $(P, -b_1, \dots, -b_n)$, where $b_1 \geq \dots \geq b_n > 0$ and $\sum b_j = P$, [2] proves

$$\kappa_d(F) = \sum_{j=1}^n j b_j, \quad (5)$$

independently of the number of ambient zeros. Here is the argument so that the sharpness and stability conclusions below do not require an omitted proof. At a minimizing common nonnegative spectrum $s_1 \geq \dots \geq s_d \geq 0$, the last coordinate is zero: otherwise subtracting $s_d I$ from both isospectral matrices preserves their difference and reduces their trace. In Horn's inequality for $(s, -s^{\text{rev}}, \lambda(F))$, choose

$$I_\ell = \{1, \dots, \ell\}, \quad J_\ell = K_\ell = \{1, d - \ell + 2, \dots, d\}, \quad 1 \leq \ell \leq n. \quad (6)$$

These triples are admissible because the partition associated with I_ℓ is zero and the other two partitions coincide; the Littlewood--Richardson coefficient is one. With the inequality written as the sum of the selected entries of the two input spectra bounded below by the selected output sum, the input sum telescopes to $s_\ell - s_d = s_\ell$, and the output sum is $P - \sum_{j < \ell} b_j$. Therefore $s_\ell \geq \sum_{j \geq \ell} b_j$, and summing gives the stated cost as a lower bound. In the eigenbasis ordered $(P, -b_1, \dots, -b_n)$, the shift

$$C e_{\ell+1} = \sqrt{2 \sum_{j \geq \ell} b_j} e_\ell \quad (7)$$

has half self-commutator equal to the target and exactly that cost; extend it by zero on the remaining coordinates. This proves the formula. Equality at the fixed-mass upper endpoint occurs precisely when the b_j are equal, as also follows from the Gini identity proved below. This lemma and its proof mechanism are recalled from [1,2].

3. Flat-block construction and arithmetic

Lemma 1 (flat-block witnesses). For integers $k, l \geq 1$, let $E_{k,l}$ have k eigenvalues $1/k$, l eigenvalues $-1/l$, and any additional zeros. Thus its positive mass is 1. Put $g = \gcd(k, l)$ and $L = (k + l)/g$. Then

$$K_d(E_{k,l}) \leq c_{kl} := \frac{(k+l)(k+l-g)}{2kl}. \quad (8)$$

Proof. Consider the modular orbit $r_j \equiv j(l/g) \pmod{L}$, represented in $\{0, \dots, L-1\}$, with $r_0 = r_L = 0$. Since $\gcd(l/g, L) = 1$, it visits every residue once before returning. Its successive differences are either l/g or $-k/g$, occurring k/g and l/g times respectively. Multiply the orbit by $g/(kl)$. Its increments are $1/k$ or $-1/l$, and all partial sums are nonnegative. A finite weighted shift whose squared weights are twice these partial sums realizes those diagonal increments as half its self-commutator. Take g disjoint copies and pad by zero. There are exactly k positive and l negative increments, and the half-squared Hilbert--Schmidt cost is

$$g \frac{g L(L-1)}{2kl} = c_{kl}. \quad (9)$$

This is a constructive upper bound, not a claim that this shift is optimal for a general two-level spectrum.

Lemma 2 (arithmetic comparison). If $\min(k, l) = 1$, then $c_{kl} = (\max(k, l) + 1)/2$. If $k, l \geq 2$, then

$$c_{kl} < \frac{\max(k, l) + 1}{2}. \quad (10)$$

Proof. Assume $l \geq k$ and write $l = k + t$, $t \geq 0$. Replacing g by 1 can only increase c_{kl} . The difference of the two numerators after multiplying by $2kl$ is

$$kl(l + 1) - (k + l)(k + l - 1) = (k - 1)\{k(k - 2) + t(2k - 1) + t^2\}. \quad (11)$$

It is positive except at $k = l = 2$. In that case $g = 2$ and $c_{22} = 1 < 3/2$, proving strictness as well.

4. Proof of the ceiling, strictness and sharpness

By homogeneity set $P = 1$. Let $a_1 \geq \dots \geq a_m > 0$ be the positive eigenvalues and $b_1 \geq \dots \geq b_n > 0$ the negative magnitudes, each list summing to 1. Define $a_{m+1} = b_{n+1} = 0$ and

$$\alpha_k = k(a_k - a_{k+1}), \quad \beta_l = l(b_l - b_{l+1}). \quad (12)$$

These are nonnegative probability vectors. The decreasing spectrum of F is the convex combination

$$\lambda(F) = \sum_{k=1}^m \sum_{l=1}^n \alpha_k \beta_l \lambda(E_{k,l}), \quad (13)$$

where the k positives occupy the first k positions and the l negatives the last l positions, with the remaining coordinates zero. This equality follows coordinate by coordinate from $\sum_{k \geq i} \alpha_k / k = a_i$ and its negative analogue. All summands and their combination lie in the same ordered chamber. Convexity and Lemmas 1--2 give the more informative certificate

$$\frac{\kappa_d(F)}{P} \leq \sum_{k,l} \alpha_k \beta_l c_{kl} \leq \frac{M+1}{2}. \quad (14)$$

When $m, n \geq 2$, the coefficient $\alpha_m \beta_n = mna_m b_n$ is strictly positive. Lemma 2 is strict at the (m, n) summand, and every other summand is bounded by the same ceiling. Thus the ceiling is strict. For normalized sign lists, restoring the mass gives the explicit positive gap

$$\frac{M+1}{2}P - \kappa_d(F) \geq Pmna_m b_n \left(\frac{M+1}{2} - c_{mn} \right) > 0. \quad (15)$$

Suppose $n = M \geq m$. For $0 < \varepsilon < 1/m$, choose positive eigenvalues $1 - (m-1)\varepsilon, \varepsilon, \dots, \varepsilon$, negative eigenvalues all $-1/n$, and the prescribed zero padding. These targets stay in the exact (m, n) stratum and converge to the zero-padded one-spike target of cost $(n+1)/2$, by the already established one-spike formula.

The cost is lower semicontinuous under this limit: if a subsequence of costs has finite $\liminf$, its minimizers have bounded Hilbert--Schmidt norm; choose a convergent subsequence, pass to the polynomial commutator constraint, and compare the limit factor

with the minimum for the limit target. Thus the limiting inferior is at least $(n+1)/2$. The ceiling supplies the reverse inequality. If $m \geq n$, apply sign reversal. This proves the sharp supremum, including arbitrary ambient zero multiplicity, without assuming ambient zero-padding invariance for general spectra (which fails in examples from [3]).

If $\min(m, n) = 1$, the existing formula $\kappa_d(F) = \sum_{j=1}^M j b_j$ for decreasing negative magnitudes (or after sign reversal) proves equality exactly at a flat opposite-sign block. The lower bound $P \leq \kappa_d(F)$ follows by the positive spectral projection trace identity in the predecessor, or directly from $\|CC^* - C^*C\|_1 \leq 2\|C\|_{HS}^2$.

5. Optimal quantitative stability

The decreasing normalized sign lists $a = (a_1, \dots, a_m)$ and $b = (b_1, \dots, b_n)$ each have mass one. Put $e_r = (1, 0, \dots, 0)$, $u_r = (1/r, \dots, 1/r)$, and

$$H = (\max(m, n) + 1)/2, \quad \delta = H - \kappa_d(F)/P, \quad h_{kl} = H - c_{kl}. \quad (16)$$

In boundary arguments, H retains the value of the fixed-inertia stratum under discussion; it is not recomputed from the smaller inertia of a limiting spectrum.

Theorem 2 (optimal constants for unequal multiplicities). For $1 \leq m < n$ and $d \geq m + n$, define $D = \|a - e_m\|_1 + \|b - u_n\|_1$. Then

$$\frac{n-m}{4(2-1/m-1/n)}D \leq \delta \leq \frac{n-1}{2}D. \quad (17)$$

Both coefficients are best possible on every stratum with exactly m positive, n negative and $d - m - n$ zero eigenvalues. For $m > n$, interchange the sign lists.

Theorem 3 (balanced stability). For $m = n = N \geq 2$, define

$$D_* = \min \{ \|a - e_N\|_1 + \|b - u_N\|_1, \|a - u_N\|_1 + \|b - e_N\|_1 \}. \quad (18)$$

With the positive finite constant

$$\gamma_N = \min_{(k,l) \notin \{(1,N), (N,1)\}} \left(\frac{N+1}{2} - c_{kl} \right), \quad (19)$$

where $1 \leq k, l \leq N$, one has

$$\frac{\gamma_N}{4(1-1/N)}D_* \leq \delta \leq \frac{N-1}{2}D_*. \quad (20)$$

The right coefficient is optimal for every N and every fixed ambient-zero multiplicity. Theorem 4 below replaces the left coefficient by its optimal value when $N = 3$; for $N = 2$ there is the exact identity $\delta = D_*/2$. For $N = 1$, both quantities vanish. All distances here are between ordered spectra.

5.1. A stronger Horn lower bound

Work at $P = 1$. For a realizing factor, put $R = CC^*/2$, $S = C^*C/2$, with common decreasing nonnegative spectrum S . The inequalities $F \leq R$ and $-F \leq S$, together with the min-max principle, give

$$s_j \geq a_j, \quad s_j \geq b_j, \quad \kappa_d(F) \geq \sum_j \max(a_j, b_j), \quad (21)$$

where shorter magnitude lists are padded by zero. At an optimum $s_d = 0$. The admissible Horn family in Section 2 applies to arbitrary sign multiplicities, and gives

$$s_l \geq a_1 - \sum_{j < l} b_j \quad (1 \leq l \leq n). \quad (22)$$

Use this inequality only for $l < n$, and replace the final term by $s_n \geq b_n$. Summing and discarding the nonnegative remaining coordinates yields

$$\kappa_d(F) \geq \sum_{j=1}^n j b_j - (n-1)(1-a_1). \quad (23)$$

For $m \leq n$, so that $H = (n+1)/2$, centering the ordered moment at u_n and using coefficients $(n+1-2j)/2$ gives

$$\delta \leq \frac{n-1}{2} \|b - u_n\|_1 + \frac{n-1}{2} \|a - e_m\|_1. \quad (24)$$

This proves the upper deficit estimate in Theorem 2. In the balanced case apply it to both signs and take the smaller distance. Individual Horn right sides may be negative; the summation argument does not require otherwise.

5.2. Closing the unbalanced atom maximum

For $m < n$, put $d_{kl} = 2(2 - 1/k - l/n)$. We claim

$$d_{kl} \leq C_{mn} h_{kl}, \quad C_{mn} = \frac{4(2 - 1/m - 1/n)}{n-m}. \quad (25)$$

For $k = 1$, direct substitution gives $d_{1l} = (4/n)h_{1l}$, including the zero-zero cell $l = n$. The positive constants C_{kn} increase with $k < n$, because their numerator increases and denominator decreases. For $k \geq 2$ it therefore suffices to use $m = k$.

Replacing $\gcd(k, l)$ by 1 increases the atom cost and can only make the claim harder. Write $q = 2 - 1/k - 1/n$ and $c_{kl}^1 = (k+l)(k+l-1)/(2kl)$. Direct algebra gives

$$2q((n+1)/2 - c_{kl}^1) - (n-k)(2 - 1/k - l/n) = -\frac{(k-1)(l-1)}{k^2 l n} B_{kn}(l), \quad (26)$$

where

$$B_{kn}(l) = k^2 l - 2k^2 n + k^2 - k l n + k l + k n + l n. \quad (27)$$

This polynomial is affine in l . Its endpoint values are

$$B_{kn}(1) = -(2k^2 - 1)(n - 1) + k + 1 < 0, \quad (28)$$

$$B_{kn}(n) = -(k + n)(kn - k - n) < 0. \quad (29)$$

Indeed $k \geq 2$ and $n \geq k + 1$: the first value is at most $-2k^3 + 2k + 1 < 0$, and $kn - k - n = (k - 1)(n - 1) - 1 \geq 1$. The claimed atom inequality follows. Equality holds at $(k, l) = (m, 1)$; when $m = 1$, the direct identity already gives the stated coefficient.

Convexity of the distance and the flat decomposition now imply

$$D \leq \sum_{k,l} \alpha_k \beta_l d_{kl} \leq C_{mn} \sum_{k,l} \alpha_k \beta_l h_{kl} \leq C_{mn} \delta. \quad (30)$$

5.3. The general balanced bound and all near extremizers

In the balanced case the only zero-gap cells are $(1, N)$ and $(N, 1)$. Set $x = \alpha_1$, $y = \alpha_N$, $u = \beta_1$, $v = \beta_N$, $w = xv + yu$, and $S = \max(x + v, y + u)$. The product decomposition gives $\delta \geq \gamma_N(1 - w)$ as a lower bound. Every flat atom has distance at most $2(1 - 1/N)$ from each reference endpoint, so

$$D_* \leq 2(1 - 1/N)(2 - S). \quad (31)$$

Because $x + y \leq 1$ and $u + v \leq 1$,

$$4w \leq (x + v)^2 + (y + u)^2 \leq S(x + y + u + v) \leq 2S. \quad (32)$$

Thus $2 - S \leq 2(1 - w)$, proving Theorem 3's lower deficit bound.

Corollary 3.1 (all nearly extremal sequences). Fix the inertia and the number of ambient zeros. The ratios $\kappa_d(F_j)/P_j$ converge to H if and only if the normalized ordered spectra approach the one-spike, opposite-flat configuration for unequal multiplicities, or the union of the two orientations for balanced multiplicities. The orientation may alternate. Both implications follow from the two positive distance bounds. This describes every possible normalized boundary limit, including sequences with varying mass.

5.4. Sharpness on every ambient-zero stratum

Normalize $P = 1$ throughout this subsection. The case $m > n$ follows by sign reversal.

For the lower coefficient of Theorem 2, take the boundary lists $a = u_m$, $b = e_n$. The opposite one-spike formula gives cost $(m + 1)/2$, deficit $(n - m)/2$, and distance $2(2 - 1/m - 1/n)$. Replace b by $(1 - \eta)e_n + \eta u_n$, with $0 < \eta < 1$. The exact prescribed inertia holds, and fixed-dimensional continuity from Section 2 makes the ratio converge to the asserted coefficient. This works for every fixed zero padding, including $m = 1$.

For the upper deficit coefficient when $2 \leq m \leq n$, let $0 < \varepsilon \leq 1/n$ and take

$$a = (1 - \varepsilon, \varepsilon, 0, \dots, 0), \quad b = u_n. \quad (33)$$

The exact cost, for every ambient dimension containing this spectrum, is

$$\kappa_d(F) = \frac{n+1}{2} - (n-1)\varepsilon. \quad (34)$$

To prove the matching upper bound, interpolate in the ordered chamber between $E_{1,n}$ and G with positive list $((n-1)/n, 1/n)$ and negative list u_n . The latter is the orthogonal sum of a paired block of mass $1/n$ and a one-spike block with $n-1$ negative magnitudes $1/n$, so its factor costs $1/n + (n-1)/2$. Section 5.1 gives the matching lower bound at G and at every point of the interpolation $(1-n\varepsilon)E_{1,n} + n\varepsilon G$.

For $m=2$, the distance is 2ε and the upper deficit coefficient is attained. For $m>2$, fix $\varepsilon = 1/(2n)$, put $a_3 = \dots = a_m = \eta$, and replace a_2 by $\varepsilon - (m-2)\eta$. For $0 < \eta < \varepsilon/(m-1)$ the list is ordered with the required inertia. Its distance stays 2ε and its cost converges to the displayed value. The same perturbation proves sharpness for balanced N : at the limiting spectrum the other orientation has distance at least $2(1-1/N)$, so the smaller distance is 2ε .

For $m=1$, take $b = u_n + t(e_1 - e_n)$ with $0 < t < 1/n$. The recalled one-spike formula gives deficit $(n-1)t$ and distance $2t$. This is the classical sharp Gini example from [2]. Sharpness may therefore be attained or approached at a boundary, as specified by these families.

5.5. The optimal balanced three-level constant

Theorem 4 (complete balanced three-level stability). For exactly three positive and three negative eigenvalues, and any fixed $d \geq 6$,

$$\frac{17}{36}D^* \leq \delta \leq D^*. \quad (35)$$

Both coefficients are optimal for every prescribed number of ambient zeros.

Proof of the lower bound. Normalize $P = 1$. We first allow nonnegative ordered lists a, b of length three and mass one, so that boundary points remain in a common compact domain. Keep $H = 2$ on this closed domain, including boundary points whose actual inertia is smaller; the extended deficit is $2 - \kappa_d$. Set $E(a) = 2(1 - a_1)$ and $U(a) = \|a - u_3\|_1$. The ordered simplex has vertices

$$e = (1, 0, 0), \quad p = (1/2, 1/2, 0), \quad u = (1/3, 1/3, 1/3). \quad (36)$$

The cut $a_2 = 1/3$ divides it into the triangles $\text{conv}\{e, v, u\}$ and $\text{conv}\{p, v, u\}$, where $v = (2/3, 1/3, 0)$. On each triangle U is affine: it is $2(a_1 - 1/3)$ when $a_2 \leq 1/3$, and $2(1/3 - a_3)$ when $a_2 \geq 1/3$.

Put $q = E - U$. On each product of two such triangles, the two orientations of the distance are separated by $q(a) = q(b)$. This gives eight polytopes on each of which D^* is affine. The values

$$q(e) = -4/3, \quad q(v) = 0, \quad q(p) = 1/3, \quad q(u) = 4/3 \quad (37)$$

make their vertices explicit. A hyperplane cut introduces only intersections with edges in addition to the surviving vertices. Every edge of a product of triangles fixes one factor at a vertex and moves the other along an edge. The union of all cut-polytope vertices is precisely the 16 ordered pairs from $\{e, v, p, u\}$, together with the following three pairs and their reversals:

$$(w, v), \quad (r, p), \quad (t, p), \quad (38)$$

$$w = (2/3, 1/6, 1/6), \quad r = (7/12, 5/24, 5/24), \quad t = (7/12, 1/3, 1/12). \quad (39)$$

For example, w and r are the points on the edge $[e, u]$ with $q = 0$ and $q = 1/3$; t is the point on $[v, u]$ with $q = 1/3$. This lists all 22 vertices. The exact replay independently enumerates all four-constraint intersections of the eight inequality descriptions and recovers the same list.

At each vertex Table 1 gives an ordered common spectrum s with zero coordinates suppressed. Pad it to length six. For the decreasing target $\lambda = (a_1, a_2, a_3, -b_3, -b_2, -b_1)$, every table entry satisfies all Horn inequalities

$$\sum_{i \in I} s_i - \sum_{j \in J} s_{7-j} \geq \sum_{k \in K} \lambda_k. \quad (40)$$

There are 522 recursive Horn triples in dimension six [11], distributed as 21, 126, 228, 126, 21 over subset sizes one through five. The exact certificate supplies and verifies every rational inequality for all 22 ordered vertices; the shorter table uses sign symmetry. Trace and ordering are checked as well. Horn sufficiency gives $\kappa_6 \leq \sum s_i$, and zero extension gives the same upper bound for every $d \geq 6$.

Table 1. Rational upper witnesses at the balanced three-level vertices. Reversing a pair preserves the common spectrum and the distance. Each row lists only nonzero entries of S ; these are upper witnesses, not a claim of optimality at every vertex.

(a, b)	Nonzero common spectrum S	D^*	$\sum_i s_i$
(e, e)	(1)	$\frac{4}{3}$	1
(e, v)	$(1, \frac{1}{3})$	$\frac{2}{3}$	$\frac{4}{3}$
(e, p)	$(1, \frac{1}{2})$	$\frac{2}{3}$	$\frac{3}{2}$
(e, u)	$(1, \frac{2}{3}, \frac{1}{3})$	0	2
(v, v)	$(\frac{2}{3}, \frac{1}{3})$	$\frac{4}{3}$	1
(v, p)	$(\frac{2}{3}, \frac{1}{2})$	$\frac{4}{3}$	$\frac{7}{6}$
(v, u)	$(\frac{2}{3}, \frac{1}{3}, \frac{1}{3})$	$\frac{2}{3}$	$\frac{4}{3}$
(p, p)	$(\frac{1}{2}, \frac{1}{2})$	$\frac{5}{3}$	1
(p, u)	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{6})$	1	$\frac{3}{2}$
(u, u)	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$	$\frac{4}{3}$	1
(w, v)	$(\frac{2}{3}, \frac{1}{3}, \frac{1}{6})$	$\frac{4}{3}$	$\frac{7}{6}$
(r, p)	$(\frac{7}{12}, \frac{1}{2}, \frac{5}{24})$	$\frac{3}{2}$	$\frac{31}{24}$
(t, p)	$(\frac{7}{12}, \frac{1}{2}, \frac{1}{12})$	$\frac{3}{2}$	$\frac{7}{6}$

For every table row,

$$\sum_i s_i + \frac{17}{36} D^* \leq 2. \quad (41)$$

On each of the eight polytopes, $\kappa_d + (17/36)D^*$ is convex because the cost is convex on the ordered chamber and the distance is affine there. Its value at every point is at most the largest vertex value, namely 2. This proves the lower deficit bound on the whole domain, and hence on the exact inertia stratum.

Sharpness. At $(a, b) = (r, p)$, the elementary lower bound in Section 5.1 gives

$$\kappa_d \geq \frac{7}{12} + \frac{1}{2} + \frac{5}{24} = \frac{31}{24}. \quad (42)$$

Table 1 gives the matching upper bound for every $d \geq 6$. Both orientations have distance $3/2$, so $\delta = 17/24$ and $\delta/D^* = 17/36$. To stay in exact inertia $(3, 3)$, replace p by $(1/2 - \eta/2, 1/2 - \eta/2, \eta)$, with $0 < \eta < 5/24$. The two orientations have distances $3/2 - 2\eta$ and $3/2 + \eta$. The elementary lower bound and the inequality just proved give

$$\frac{31}{24} - \frac{\eta}{2} \leq \kappa_d \leq \frac{31}{24} + \frac{17\eta}{18}. \quad (43)$$

Thus the cost tends to $31/24$ and the deficit-to-distance ratio tends to $17/36$, independently of the fixed zero multiplicity. This explicit sandwich also avoids relying on continuity for the new sharpness result. The upper coefficient was proved sharp in Section 5.4, completing the theorem.

For comparison, balanced multiplicity two satisfies $\kappa_d/P = 1 + |a_1 - b_1|$ for every $d \geq 4$: the elementary lower bound matches the two-by-two transport construction in Section 6. This is the known four-level formula [9], with zero extension justified by the same matching bounds. Direct evaluation gives $D^* = 1 - 2|a_1 - b_1|$, hence $\delta = D^*/2$. The coefficient $17/36$ for multiplicity three is different; a universal extrapolation of the multiplicity-two coefficient would be false.

6. Optimizing the flat-block coupling

Continue to use normalized sign lists and their probability coefficients α, β . Define the transportation polytope and its cost by

$$\Pi(\alpha, \beta) = \{\pi \geq 0 : \sum_l \pi_{kl} = \alpha_k, \sum_k \pi_{kl} = \beta_l\}, \quad T(\alpha, \beta) = \min_{\pi \in \Pi(\alpha, \beta)} \sum_{k,l} \pi_{kl} c_{kl}. \quad (44)$$

Theorem 5 (transport refinement). For every nonzero traceless Hermitian target,

$$\frac{\kappa_d(F)}{P} \leq T(\alpha, \beta) \leq \sum_{k,l} \alpha_k \beta_l c_{kl} \leq H. \quad (45)$$

An optimum exists with at most $|\text{supp } \alpha| + |\text{supp } \beta| - 1 \leq m + n - 1$ positive entries. Its finite linear-programming dual is

$$T(\alpha, \beta) = \max_{u_k + v_l \leq c_{kl}} \left(\sum_k \alpha_k u_k + \sum_l \beta_l v_l \right). \quad (46)$$

The potentials u_k, v_l are unrestricted real numbers. For rational sign lists, rational primal and dual optima exist. Matching primal and dual objectives certify the optimal value of this upper-bound construction; a transport dual is not a lower bound on the matrix cost.

Proof. Every coupling reconstructs the same ordered target, because its positive coordinate i is $\sum_{k \geq i} \alpha_k/k = a_i$, and its negative coordinate $d - j + 1$ is $-\sum_{l \geq j} \beta_l/l = -b_j$. All remaining coordinates vanish. Horn convexity and Lemma 1 therefore apply to every plan. The product plan is feasible, and the transportation polytope is compact and nonempty. This proves the inequalities and attainment.

The dual and equality of optimal values are standard finite linear-programming duality [8]. Directly, the difference between feasible primal and dual values is

$$\sum_{k,l} \pi_{kl} (c_{kl} - u_k - v_l) \geq 0. \quad (47)$$

It vanishes when every positive entry of π is dual-tight. To prove sparsity, remove zero marginal rows and columns and choose an optimal extreme point. A cycle in its positive-entry bipartite graph would allow small alternating perturbations of both signs,

preserving the margins and nonnegativity, contradicting extremality. A forest on $r + s$ active vertices has at most $r + s - 1$ edges. Rationality follows from the rational coefficients of the two finite programs.

This is the strongest pointwise bound obtainable from convexity and these particular flat-atom costs alone. Indeed, the decreasing nonnegative mass-one lists form a simplex whose vertices are the padded flat lists. Their barycentric coordinates α and β are unique. Thus every convex decomposition over the vertices of the product of the two simplices is exactly a transportation plan. The function T is convex by mixing optimal plans, equals C_{kl} at each vertex, and majorizes any convex function whose vertex values are bounded by C_{kl} , by Jensen's inequality. This statement concerns the envelope of the specified vertex data; it does not say that the chosen flat witnesses or T equal the exact inverse cost.

Proposition 5.1 (when the product can improve). On the active rectangle $\text{supp } \alpha \times \text{supp } \beta$, the product plan is optimal if and only if C_{kl} is the sum of a row term and a column term, equivalently all its 2-by-2 contrasts vanish. If a contrast is nonzero, an alternating rectangular perturbation of the strictly positive product plan decreases its cost. If all contrasts vanish, fixing a row and column gives the additive form, and every plan has the same cost. This also covers singleton supports.

In particular, $c_{11} = c_{22} = 1$ and $c_{12} = c_{21} = 3/2$. Moving the smaller cross mass into the two diagonal cells yields

$$T(\alpha, \beta) \leq \sum_{k,l} \alpha_k \beta_l c_{kl} - \min(\alpha_1 \beta_2, \alpha_2 \beta_1) \quad (48)$$

when $m, n \geq 2$. For $a = b = (3/4, 1/4)$, the product bound is $5/4$, whereas the plan $\pi_{11} = \pi_{22} = 1/2$ costs 1. The dual $u = (0, 0)$, $v = (1, 1)$ proves $T = 1$. Two disjoint paired shifts attain the general floor $\kappa_d(F)/P = 1$ here; this paired endpoint was already known [1]. For the entire 2-by-2 margin family $\alpha = (x, 1 - x)$, $\beta = (y, 1 - y)$,

$$T = 1 + \frac{1}{2}|x - y|. \quad (49)$$

If $x \geq y$, use $\pi_{11} = y$, $\pi_{12} = x - y$, $\pi_{22} = 1 - x$ and dual $u = (1/2, 0)$, $v = (1/2, 1)$. If $x \leq y$, use $\pi_{11} = x$, $\pi_{21} = y - x$, $\pi_{22} = 1 - y$ and dual $u = (0, 1/2)$, $v = (1, 1/2)$. All unlisted plan entries are zero. These certificates include the boundary cases.

An additional exact example has $a = (3/4, 1/4)$ and $b = (5/12, 5/12, 1/6)$. The product bound is $37/24$; the plan $\pi_{13} = \pi_{22} = 1/2$ costs $3/2$, certified by $u = (0, -1/2)$ and $v = (1, 3/2, 2)$. Its exact normalized matrix cost is $4/3$ for every $d \geq 5$: the common spectrum $(3/4, 5/12, 1/6, 0, 0)$ satisfies all 142 Horn inequalities in dimension five, while Section 5.1 gives the matching ambient-independent lower bound $3/4 + 5/12 + 1/6$. Horn sufficiency and zero extension prove the upper bound in every larger dimension. This benchmark is derived from the existing five-level theory [10]; it is not a new five-dimensional cost formula. Thus the optimized transport value exceeds the actual cost by $1/6$, or 12.5%. The accompanying exact replay includes the full certificate. The cost contrasts on rows (1, 2) and columns (1, 2) or (2, 3) are respectively -1 and $1/6$, so a universal monotone-coupling rule is unjustified. Zero marginal entries are removed without renumbering the remaining

atom indices. An optimal plan can have $\pi_{mn} = 0$ even when both final marginals are positive; the explicit gap in Section 4 and the stability proof retain the product plan.

Sparsity bounds the number of atoms in a spectral certificate, not the rank of a factor. Adding the flat factors themselves would introduce cross terms; synthesis in dimension d still uses Horn sufficiency. The transport duality, envelope description and sparsity are classical optimization mechanisms [8], applied here to the arithmetic flat-block costs.

7. Stability near the rank-only ceiling

Corollary 5.2 (sharp sign-purity threshold). Let $\rho \geq 3$ and define

$$\Delta = \frac{\rho}{2} - \frac{\kappa_d(F)}{P}. \quad (50)$$

If $\Delta \leq 1/2$, then $\min(m, n) = 1$. Indeed, if both are at least two, strictness gives $\Delta > (\min(m, n) - 1)/2 \geq 1/2$. After changing sign, write the normalized spectrum as one positive eigenvalue 1, $N = \rho - 1$ negative magnitudes $u_1 \geq \dots \geq u_N > 0$, summing to 1, and zeros. Put $\mathcal{L} = \sum_j |u_j - 1/N|$. Then the exact Gini identity from [2] yields

$$\frac{N}{4}\mathcal{L} \leq \Delta \leq \frac{N-1}{2}\mathcal{L}, \quad \mathcal{L} \leq \frac{4\Delta}{\rho-1}. \quad (51)$$

For completeness, $2\Delta = \sum_{i < j} (u_i - u_j)$. Subtract $1/N$, split at the sign of this deviation, and keep cross pairs: their sum is $N\mathcal{L}/2$. This gives the lower inequality. The coefficients of the centered linear functional are $(N + 1 - 2j)/2$, of absolute value at most $(N - 1)/2$, giving the upper inequality. Two-level centered deviations attain the lower constant; deviations at only the first and last coordinates attain the upper constant. They can be arbitrarily small, so the constants remain sharp in the stated neighborhood.

For every $\rho \geq 4$, the threshold $1/2$ cannot be increased: the sharpness family with $(m, n) = (2, \rho - 2)$ has both signs multiple and deficit tending down to $1/2$. This closes the local arbitrary-sign stability problem left explicit in [1]. The Gini constants themselves were already proved in [2].

8. Verification, antecedents and remaining questions

The main result controls norm cost; it does not bound the rank of norm-optimal factors by the larger sign multiplicity [3]. The optimal stability constants are now determined in both directions for every unequal pair of sign multiplicities and for balanced multiplicities two and three. The reverse coefficient is optimal for every balanced multiplicity; the optimal forward coefficient for $N \geq 4$ remains open. The general interior cost and the classification of optimizing matrices are not determined here. General zero-padding invariance is not assumed.

The new balanced-three result has a finite, exact computational proof. The domain decomposition and its completeness are justified in Section 5.5. The replay uses rational arithmetic to enumerate 2,640 active sets, recover eight cell vertex sets and 22 distinct vertices, generate the 522 Horn triples, and check 11,484 Horn inequalities together with the stability inequality and the matching sharpness lower bound. Floating-point

optimization proposed the common spectra during exploration; every proposal used in the proof was subsequently verified exactly. The imported Horn sufficiency theorem is a classical input [4,5,11], not a new result or a proof-assistant formalization.

The consolidated review package also supplies the earlier ceiling, transportation and optimal-constant replay programs. These check weighted shifts, spectral reconstructions, rational primal-dual equality, the polynomial identities in Section 5.2, the explicit sharpness families and the 142-inequality transport separation benchmark. Integer-grid checks support implementation correctness; all-multiplicity validity rests on the algebraic proofs. Separate internal review checks the arguments and regenerates the Horn triples by Littlewood-Richardson tableaux rather than the recursive generator.

Relative to the inspected antecedents, the contributions are the sharp fixed-inertia ceiling, optimal unbalanced stability, the optimal reverse coefficient for all inertias, the optimal balanced-three forward coefficient, the classification of all near extremizers, and the certified transport refinement with its demonstrated limitation. The one-spike formula [2], the four- and five-level formulas [9,10], weighted shifts [6,7], and transport duality [8] retain their original attribution. This version absorbs the two earlier private manuscripts, rather than presenting their overlapping results as independent publications.

A bounded primary-source comparison did not identify an external statement of these specific constants or the inertia ceiling. That search does not establish priority. The finite eigenvalue-sum criterion and its recursive indexing are described by Fulton [11]. The review package records the actual scope of the computations and internal review.

AI assistance and review status. Lluís Eriksson directed the research programme. OpenAI Codex assisted with proof development, exact computation, source comparison, separate internal critical review and drafting. Internal AI review is not independent human refereeing. This consolidated manuscript is supplied for the author's feedback; it has not been published as this revision.

References

- [1] Lluís Eriksson. *Sharp Rank-Adaptive Bounds for Inverse Self-Commutators*. ARR-2026-1D2QV1RP1292JREW, v1 (2026). [Record](#).
- [2] Lluís Eriksson. *One-Spike Inverse Self-Commutators and Exact Three-versus-Four-Kick Curvature Synthesis*. ARR-2026-7NPRNBW4488HG90K, v1 (2026). [Record](#).
- [3] Lluís Eriksson. *Sharp Onset and Unbounded Growth of Norm-Optimal Self-Commutator Rank*. ARR-2026-5QQF95VHTC9GABH8, v1 (2026). [Record](#).
- [4] A. A. Klyachko. *Stable bundles, representation theory and Hermitian operators*. Selecta Mathematica 4 (1998), 419–445. [DOI](#).
- [5] A. Knutson and T. Tao. *The honeycomb model of $GL(n)$ tensor products I: Proof of the saturation conjecture*. Journal of the American Mathematical Society 12 (1999), 1055–1090. [DOI](#).
- [6] P. Fan and C. K. Fong. *Which operators are the self-commutators of compact operators?* Proceedings of the American Mathematical Society 80 (1980), 58–60. [DOI](#).
- [7] D. Beldiș, S. Patnaik and G. Weiss. *$B(H)$ -Commutators: A Historical Survey II and recent advances on commutators of compact operators*. [arXiv:1303.4844](#), especially Section 4.
- [8] G. Peyré and M. Cuturi. *Computational Optimal Transport*. Foundations and Trends in Machine Learning 11 (2019), 355–607; Sections 2.5 and 3.3–3.4. [Author version](#).
- [9] Lluís Eriksson. *The Exact Four-Level Inverse Commutator Cost*. ARR-2026-3M1EEG1T689ADSMW, v1 (2026). [Record](#).
- [10] Lluís Eriksson. *The Exact Five-Level Inverse Commutator Cost*. ARR-2026-37B8R0QTA894GTFF, v1 (2026). [Record](#).
- [11] William Fulton. *Eigenvalues, invariant factors, highest weights, and Schubert calculus*. Bulletin of the American Mathematical Society 37 (2000), 209–249. [arXiv:math/9908012](#).