hexagon:2611.01281v1

QUANTITATIVE GAUSSIAN SIMPLEX STABILITY AT SMALL POSITIVE CORRELATIONS

ROY VAN RIJN

Abstract. For each fixed number of cells $k \geq 2$, we prove that the partition defined by a centered regular simplex maximizes Gaussian noise stability among equal-measure measurable k-cell partitions at sufficiently small positive correlations ρ in dimensions $d \geq k-1$. The noise stability deficit is bounded below by a positive constant times ρ times the square of the summed Gaussian symmetric differences from the closest simplex partition. The constant and the correlation threshold are independent of the ambient dimension. Zhang's first-moment theorem localizes near maximizers. To prove the positive-correlation bound, we normalize arbitrary label changes by their exact loss against the simplex's Ornstein–Uhlenbeck scores. A Gaussian covariance estimate on the facets bounds the limiting first-moment interaction, and a Hermite expansion controls all remaining directions uniformly in the dimension.

1. The main result

Partition Gaussian space into k cells of equal measure. When two Gaussian observations have positive correlation, which partition makes them most likely to have the same label? The Gaussian Standard Simplex Conjecture predicts that the cells should be the winning regions of a centered regular simplex [6, Conjecture 1.4(i)]. We prove a quantitative form of the comparison near zero correlation.

Write γ_d for standard Gaussian probability on $\mathbb{R}^d$, with density ϕ_d. For $0 < \rho < 1$ let

$$T_\rho f(x) = \int f(\rho x + \sqrt{1-\rho^2}\, y)\, d\gamma_d(y), \qquad S_\rho(A) = \sum_i \int 1_{A_i} T_\rho 1_{A_i}\, d\gamma_d.$$

Thus $S_\rho(A)$ is the probability of agreement. A measurable partition $A = (A_1, \ldots, A_k)$ is *balanced* when $\gamma_d(A_i) = 1/k$ for each i. For $n = k-1$ and $d \geq n$, take the vertices v_i of a centered regular simplex in an n-dimensional subspace and define

$$T_i = \{x : v_i \cdot x \geq v_j \cdot x \text{ for every } j\}.$$

For $k = 3$ these are three 120° sectors, with any remaining coordinates unrestricted. We identify partitions up to Gaussian null sets. The *simplex orbit* consists of all such labeled partitions in $\mathbb{R}^d$. Let

$$d_k(A) = \inf_{T' \text{ in the simplex orbit}} \sum_i \gamma_d(A_i \triangle T'_i).$$

For fixed T', this sum is twice the probability of different labels.

Theorem 1.1. *For every integer $k \geq 2$ there are constants $c_k > 0$ and $0 < \rho_k < 1$ such that in every dimension $d \geq k-1$ and for every balanced measurable k-cell partition A,*

$$S_\rho(T) - S_\rho(A) \;\geq\; c_k \rho\, d_k(A)^2 \qquad (0 < \rho \leq \rho_k).$$

Equality of noise stabilities holds exactly on the simplex orbit.

2020 *Mathematics Subject Classification.* 60G15, 49K21, 28A75.

Key words and phrases. Gaussian noise stability, regular simplex, standard simplex conjecture, first moments.

The constants c_k and ρ_k may depend on k and are left unevaluated. The quantitative conclusion is the quadratic control of label distance. If $d < k-1$, cylindrical extension to dimension $k-1$ gives the strict upper bound $S_\rho(A) < S_\rho(T)$ for $0 < \rho \le \rho_k$, with T in dimension $k-1$. The extended partition's cell moments cannot span that space.

Borell proved two-cell simplex optimality [1]. For three cells, Heilman and Tarter extended Heilman's fixed-dimensional result to all dimensions [2, 5]. Heilman later reached $\rho = 0.1082$, with a planar radial-mass deficit [3, Theorem 1.5]. His 2026 preprint proves the balanced three-cell comparison for $0 < \rho \le 2/5$ in every $d \ge 2$ [4, Theorems 1.2 and A.1] and leaves stronger estimates of distance to the optimizer for future work [4, Section 1.4].

Theorem 1.1 applies to every fixed cell count, including $k \ge 4$. Its deficit controls the full Gaussian label distance to the simplex orbit, with both constants uniform in the ambient dimension. This is the Gaussian problem underlying low-influence Plurality Is Stablest [6].

How the comparison works. The coefficient of ρ in $S_\rho(A)$ is the sum of squared Gaussian cell moments. Zhang proved that the simplex maximizes this coefficient for every cell count [9]. We include the covariance argument for that theorem with attribution. It implies that competitors with stability at least $S_\rho(T) - o(\rho)$ approach the simplex orbit as $\rho \downarrow 0$. An absolute $O(\rho^2)$ error estimate gives too little information once the competitor is extremely close.

For the local comparison, keep the simplex scores $T_\rho 1_{T_i}$ fixed while changing labels. The exact stability change is the quadratic interaction of the changes minus twice their score loss. Across a facet, the two corresponding scores agree, and the slope of their difference is asymptotic to $\rho\ell$, where ℓ is the distance between the two cell moments. The leading cost of moving a thin strip of width h is therefore $\rho\ell h^2/2$ per unit Gaussian surface area. This gives the scale for arbitrary changes, even those with no smooth boundary. A bound on the Gaussian tails and the same strip estimate yield limiting densities on the simplex facets. Aligning the cell moments removes infinitesimal rotations. The conditional Gaussian covariance on each convex facet then bounds the limiting ratio of interaction to score loss by $2 - 2/k$ (Proposition 6.1). All higher Hermite degrees contribute a uniformly smaller error, including in directions orthogonal to the simplex. Proposition 7.1 exchanges two remote balls to explain why the first-moment deficit cannot replace the exact score loss in this local argument.

2. First moments and uniform localization

Use $E = \mathbf{1}^\perp \subset \mathbb{R}^k$, the projection $P = I - \mathbf{1}\mathbf{1}^T/k$, and the reference covariance $Q_* = P/n$. For a partition let B be the matrix with rows $b_i = \int_{A_i} x\, d\gamma_d(x)$ and put $W_1(A) = \|B\|_F^2$, where $\|\cdot\|_F$ denotes the Frobenius norm. The sum of the rows is zero. Write

$$F = \frac{(\mathbb{E}\max_{i\le k} Z_i)^2}{n},$$

where the Z_i are independent standard normals.

Proposition 2.1. *For every balanced partition, $W_1(A) \le F$. Equality holds exactly on the simplex orbit when $d \ge n$, and the inequality is strict for $d < n$ [9, Theorem 1]. Moreover, for $d \ge n$, the label distance to the simplex with moment matrix closest to B tends to zero as $W_1(A) \to F$, uniformly in d.*

Proof. We present Zhang's covariance comparison [9, Sections 2–5] with an additive parametrization. We then derive the uniform localization using compactness of the normalized covariances and a linear score estimate. For $Q \succeq 0$ with $Q\mathbf{1} = 0$ let

$$C(Q) = \inf_{\lambda\in\mathbb{R}^k} \left\{ \frac{1}{k}\sum_i \lambda_i + \mathbb{E}\max_i (Y_i - \lambda_i) \right\}, \qquad Y \sim N(0, Q).$$

It is nonnegative (the maximum dominates the average), homogeneous of degree $1/2$, and continuous even at singular Q. Coupling the scores via the same standard Gaussian bounds the change of the objective uniformly in λ by the expected norm of the score change. For Q positive on E, represent $Y = Mx$ in dimension n with $MM^T = Q$. On the gauge $\sum_i \lambda_i = 0$ the objective is at least $-\min_i \lambda_i$, hence a minimizer exists. Differentiation in the prices gives equal masses, since ties have measure zero. Each pair of winning cells has a genuine facet: the map $x \mapsto Mx$ is onto E, so each prescribed strict ordering between the tied pair and the remaining labels occurs on a relatively open portion of their tie plane. Any two minimizing price vectors give the same winners almost everywhere, by testing the balanced assignment of one vector against the objective of the other. Different price differences for a pair would shift its cutoff and change the winner on an open set where the other scores are lower. Thus the normalized minimizer is unique, and it varies continuously with M by coercivity and continuity of the objective.

Let ℓ_{ij} be the distance between score sites and A_{ij} the Gaussian area of their winning facet. Let L be the graph Laplacian on the labels with weights $w_{ij} = A_{ij}/\ell_{ij}$. Gaussian integration by parts gives the winning moment matrix LM, since the outward normal for cell i on its facet with j is $(M_j - M_i)/\ell_{ij}$. The boundary integral on a bounding sphere tends to zero. Evaluating the score at the balanced winners gives

$$C(Q) = \operatorname{tr}(LQ) = \sum_{i<j} w_{ij}\ell_{ij}^2, \qquad dC(Q)[H] = \tfrac{1}{2}\operatorname{tr}(LH).$$

For the derivative write $H = MKM^T$ with K symmetric and vary the sites by $M(I + tK)^{1/2}$. Differentiation at the unique minimizing prices gives the envelope derivative from the winning moments. Indeed, the difference of the minimum values is bounded using the minimizing price vector at each endpoint, and the two derivatives agree in the limit by continuity of the prices and nullity of ties.

At Q_* the scores have the law of $(Z_i - \overline{Z})/\sqrt{n}$, so $C(Q_*) = \sqrt{F}$. Symmetry gives $L_* = \sqrt{F}P$. Thus the regular score partition has moment Gram matrix $L_*Q_*L_* = FQ_*$, and its sum of Gaussian facet areas is $S_* = \sqrt{nF/2}$. The Gaussian multi-bubble theorem [7, Theorem 1.1] in dimension n gives $\sum_{i<j} A_{ij} \geq S_*$ for the balanced winners. The perimeter counts each interface once. For $k = 2$ the trace-one covariance is already Q_*, so the following comparison is immediate in that case.

For centered trace-one Q and $s > 0$, put $g(s) = C(Q + sQ_*)^2$. The derivative formula and weighted Cauchy–Schwarz give

$$g'(s) = C(Q + sQ_*)\frac{\operatorname{tr} L}{n} \;\geq\; \frac{2}{n}\left(\sum_{i<j} A_{ij}\right)^2 \geq F.$$

If $Q \neq Q_*$ the first inequality is strict: all weights are positive, and equal pair distances of centered sites would give a scalar multiple of P. By homogeneity and the derivative at Q_*, $g(s) = Fs + F + o(1)$ as $s \to \infty$. Hence $g(s) - Fs$ increases to F, and continuity at zero proves $C(Q)^2 \leq F$, strictly off Q_*.

Using the labels of A as a balanced assignment gives $W_1(A) \leq C(BB^T)$. When $W_1(A) > 0$, scaling gives $\sqrt{W_1(A)} \leq C(BB^T/W_1(A)) \leq \sqrt{F}$. If $W_1(A) = F$, the covariance is regular and equality in the assignment requires the original labels to be the regular winners almost everywhere. The prices are equal by symmetry and the winning labels are unique off the pair planes. The regular Gram matrix has rank n, which excludes equality for $d < n$.

For the uniform localization, the centered trace-one covariance set is compact. If $W_1(A) \to F$, the last bounds and strict optimality of Q_* force $BB^T \to FQ_*$, independently of d. Every matrix with Gram matrix FQ_* is the moment matrix of a simplex. If B_* is closest to B among these matrices, singular value decomposition gives $\|B - B_*\|_F^2 = \sum_{j=1}^n (s_j(B) - \sqrt{F/n})^2 \to 0$. Here

the singular values are padded by zeros when needed, and $d \geq n$ permits completion of the right singular vectors.

For later use, the linear score regret against the winners of B_* is exactly

$$\int (\max_i b_i^* \cdot x - b_{I_A(x)}^* \cdot x)\, d\gamma_d(x) = F - \langle B_*, B \rangle_F = \tfrac{1}{2}(F - W_1(A) + \|B - B_*\|_F^2), \tag{1}$$

where I_A is the label assigned by A. The mass where the top two regular scores differ by at most $u > 0$ is at most $C_k u$, by a union bound over nondegenerate one-dimensional Gaussian pair differences. If the integral in (1) is r, the changed-label mass is at most $C_k u + r/u$. This proves uniform localization. □

3. The aligned facet inequality

Identify the simplex span with E and write $x \in E$ for the active coordinates. If e_i are the coordinate vectors in $\mathbb{R}^k$, the regular moment rows are $b_i^* = aPe_i$ with $a = \sqrt{F/n}$. Put $p_{ij} = e_i - e_j$, $n_{ij} = p_{ij}/\sqrt{2}$, and $\ell = \sqrt{2}a$. The pair plane H_{ij} is $n_{ij}^\perp$ in E, the common facet is $F_{ij} = T_i \cap T_j$ in E, and $d\nu_{ij} = \phi_n\, d\mathcal{H}^{n-1}$ on H_{ij}, where $\mathcal{H}^{n-1}$ is Hausdorff measure. The common facet area $v = \nu_{ij}(F_{ij})$ satisfies $\ell = kv$: Gaussian flux gives $b_i^* = v \sum_{j \neq i} n_{ij} = kvPe_i/\sqrt{2}$.

For facet densities $f = (f_{ij})_{i<j}$ in $\mathcal{H} = \bigoplus_{i<j} L^2(F_{ij}, \nu_{ij})$, set $f_{ji} = -f_{ij}$. Call them balanced if $\sum_{j \neq i} \int f_{ij}\, d\nu_{ij} = 0$ for every i. Define the moment matrix $\mathcal{D}f$ by $(\mathcal{D}f)_i = \sum_{j \neq i} \int x f_{ij}\, d\nu_{ij}$. Its rows lie in E and sum to zero, so we view $\mathcal{D}f$ as an operator $E \to E$. In the local comparison, choosing the closest simplex moment matrix makes the active moment difference orthogonal to infinitesimal rotations. This leaves a symmetric operator on E, as shown in Section 6.

Lemma 3.1. *For balanced f with $\mathcal{D}f$ symmetric,*

$$\|\mathcal{D}f\|_F^2 \;\leq\; (1 - 1/k)\ell \|f\|_{\mathcal{H}}^2.$$

Proof. Put $H = \mathcal{D}f$. By symmetry under permutations preserving the unordered pair, the mean of x under ν_{ij}/v on F_{ij} is $\alpha P(e_i + e_j)$ for a constant α common to the pairs. The mean of $p_{ij}^T H x$ is therefore $\alpha(H_{ii} - H_{jj})$. These mean terms cancel when paired with the densities by balance.

The covariance of the Gaussian on the pair plane conditioned on F_{ij} is bounded by the identity on the plane. Indeed, in plane coordinates its log moment generating function is $|u|^2/2 + \log Z(u) - \log Z(0)$, where $Z(u) = \int_{F_{ij}} \phi_{n-1}(x - u)\, dx$. The Gaussian integrand with the facet indicator is jointly log concave, so Z is log concave by Prékopa's marginal theorem [8, Theorem 6]. Differentiate twice at zero. If $n = 1$ the pair plane is a point and the covariance is zero.

Writing Π_{ij} for projection onto H_{ij}, Cauchy–Schwarz after subtracting the means gives

$$\|H\|_F^2 = \sum_{i<j} \int f_{ij}\, p_{ij}^T H x\, d\nu_{ij} \leq \|f\|_{\mathcal{H}} \left(v \sum_{i<j} |\Pi_{ij} H p_{ij}|^2 \right)^{1/2}.$$

The unprojected sum is $k\|H\|_F^2$, since $\sum_{i<j} p_{ij} p_{ij}^T = kP$. The removed sum is $\frac{1}{2}\sum_{i<j}(p_{ij}^T H p_{ij})^2$. If J is the matrix with entries $J_{ij} = p_{ij}^T H p_{ij}$ and zero diagonal, then $H = -PJP/2$ by centering. The removed sum is at least $\|H\|_F^2$ because

$$\|H\|_F^2 = \tfrac{1}{4}\|PJP\|_F^2 \leq \tfrac{1}{4}\|J\|_F^2 = \tfrac{1}{2}\sum_{i<j} J_{ij}^2.$$

The projected sum is therefore at most $(k-1)\|H\|_F^2$. Squaring the Cauchy–Schwarz estimate and using $v = \ell/k$ proves the claim, with the case $H = 0$ immediate. □

4. Exact score loss and weighted transfer bounds

Put $\sigma = \sqrt{1-\rho^2}$ and $U_{\rho,ij} = T_\rho(1_{T_i} - 1_{T_j})$. In coordinates $x = z + tn_{ij}$, $z \in H_{ij}$, let $\kappa = \phi_1(0)$; then

$$d\gamma_n = \phi_1(t)dt\, d\gamma_{n-1}(z), \qquad d\nu_{ij} = \kappa\, d\gamma_{n-1}(z).$$

For $n = 1$, γ_0 is unit mass on the pair plane.

Lemma 4.1. *There is $c_T > 0$ such that for every $0 < \rho < 1$ and $x \in T_i$,*

$$U_{\rho,ij}(x) \geq c_T \min\{\rho n_{ij} \cdot x/\sigma, 1\}. \tag{2}$$

Also $U_{\rho,ij} = 0$ on H_{ij} and $\nabla U_{\rho,ij}/\rho \to b_i^* - b_j^*$ locally uniformly as $\rho \downarrow 0$.

Proof. The transposition of vertices i, j is reflection G in H_{ij}. Cone scaling, with $m = \rho x/\sigma$, gives

$$U_{\rho,ij}(x) = \int_{T_i} [\phi_n(y-m) - \phi_n(Gy-m)]dy, \qquad \frac{\phi_n(Gy-m)}{\phi_n(y-m)} = e^{-2m_N y_N}.$$

Here the subscript N denotes the n_{ij} component. For $m \in T_i$ the integrand is nonnegative. Fix a ball D in the interior of T_i on which every $u \in D$ has $u \cdot n_{ij} \geq a > 0$, and set $p = \gamma_n(D) > 0$. Since $m + D \subset T_i$, restriction to the translated ball gives $p(1 - e^{-2am_N}) \geq p(1 - e^{-2a}) \min\{m_N, 1\}$. Choose constants common to the finitely many pairs. Reflection also gives exact vanishing on the entire pair plane.

The Mehler kernel relative to $\gamma_n \otimes \gamma_n$ is

$$M_\rho(x,y) = (1-\rho^2)^{-n/2} \exp\left(\frac{2\rho x \cdot y - \rho^2(|x|^2 + |y|^2)}{2(1-\rho^2)}\right). \tag{3}$$

Its value and correlation derivative at zero are 1 and $x \cdot y$. Also $\rho^{-1}\nabla_x M_\rho = (y - \rho x)M_\rho/(1-\rho^2)$ tends to y uniformly on compact products. For x in a fixed compact set and ρ small it is bounded by a constant times $(1 + |y|)e^{|y|^2/16}$. Integrating gives the locally uniform gradient limit for the scores. □

In the active space let $0 \leq h_{ij} \leq 1$ be supported in T_i for $i \neq j$, and set $e_i = \sum_{j \neq i}(h_{ji} - h_{ij})$. For actual label changes $h_{ij} = 1_{T_i \cap A_j}$ and $e_i = 1_{A_i} - 1_{T_i}$. Define the score loss by

$$L_\rho(h) = \sum_{i \neq j} \int U_{\rho,ij} h_{ij} d\gamma_n \geq c_T \rho \sum_{i \neq j} \int \min\{n_{ij} \cdot x, 1\} h_{ij} d\gamma_n, \tag{4}$$

because $\min\{\rho t/\sigma, 1\} \geq \rho \min\{t, 1\}$. The loss of a partition is denoted by $\mathcal{L}_\rho(A;T)$, and is positive whenever $A \neq T$ modulo null sets. In dimension $d > n$ the scores still depend only on the active coordinates, so the full loss is the integral of the slice losses against the transverse Gaussian.

Lemma 4.2. *Let $w(x) = (1+|x|^2)e^{|x|^2/16}$. Uniformly over unit vectors u and measurable $0 \leq h \leq 1$ supported in $t = u \cdot x \geq 0$,*

$$\left(\int_{|x|>R} wh\, d\gamma_n\right)^2 \leq \beta(R) \int \min\{t, 1\} h\, d\gamma_n \qquad (R \geq 2), \tag{5}$$

where $\beta(R) \to 0$. There is a finite constant for the same estimate without the tail restriction.

Proof. For $t \geq 1$, Cauchy–Schwarz gives the square bound $\int_{|x|>R} w^2 d\gamma_n$ times the mass, which is its cost. For $0 \leq t < 1$ write $x = z + tu$ and $m(z) = \int_0^1 h(z + tu)\phi_1(t)dt$. Every density $v \leq \kappa$ on the positive halfline with mass m satisfies

$$\int tv(t)dt = \int_0^\infty \int_s^\infty v(t)dt\, ds \geq \int_0^{m/\kappa} (m - \kappa s)ds = \frac{m^2}{2\kappa}.$$

Thus

$$m(z)^2 \le 2\kappa \int_0^1 th(z+tu)\phi_1(t)dt. \tag{6}$$

In the tail, $|z| > \sqrt{R^2-1}$ and $w(z+tu) \le 2e^{1/16}(1+|z|^2)e^{|z|^2/16}$. Apply Cauchy–Schwarz in z and (6). The two remaining factors are the tails of w^2 against γ_n and of $(1+|z|^2)^2 e^{|z|^2/8}$ against γ_{n-1}. Both vanish. For $n=1$ the second tail is empty. Omitting tail restrictions gives the global constant. □

Lemma 4.3. *For $0 < \rho \le 1/16$, $K_\rho = (M_\rho - 1)/\rho$ satisfies*

$$|K_\rho(x,y)| \le Cw(x)w(y). \tag{7}$$

It converges to $x \cdot y$ uniformly on compact products.

Proof. Use $K_\rho = \int_0^1 \partial_r M_r|_{r=s\rho} ds$. The derivative is a quadratic polynomial with bounded coefficients times M_r, and (3) gives $M_r \le C\exp[(|x|^2+|y|^2)/34]$ on this range. The weight $w(x)w(y)$ dominates their product. The derivative at zero gives the compact limit. □

In any dimension let Π_0 be projection onto constants and write $\widetilde{Q}_\rho(e) = \sum_i \langle e_i, (T_\rho - \Pi_0)e_i\rangle$, with the inner product in the ambient Gaussian L^2 space. For balanced A and T, each e_i has mean zero. Expanding $S_\rho(A)$ about T gives

$$S_\rho(A) - S_\rho(T) = \widetilde{Q}_\rho(e) - 2\mathcal{L}_\rho(A;T). \tag{8}$$

Corollary 4.4. *For active transfers as above, uniformly for $0 < \rho \le 1/16$ and $R \ge 2$,*

$$\left(\sum_i \int_{|x|>R} w|e_i|d\gamma_n\right)^2 \le C_k\beta(R)L_\rho(h)/\rho. \tag{9}$$

The same estimate holds with no tail restriction and a finite constant. Both estimates also hold with $\sum_{i\neq j} h_{ij}$ in place of $\sum_i |e_i|$.

Proof. Sum Lemma 4.2 over ordered transfers and use (4) and $\sum_i |e_i| \le 2\sum_{i\neq j} h_{ij}$. □

5. Limits of measurable transfers

The next proposition converts arbitrary small label changes, including averages of changes in higher dimension, into densities on the facets.

Proposition 5.1. *Let $\rho_m \to 0$ through positive values and let $h^{(m)}$ be active transfers as in Section 4, with induced changes $e_i^{(m)}$ of mean zero. Suppose $s_m^2 = L_{\rho_m}(h^{(m)})/\rho_m$ is positive and tends to zero. Along a subsequence there are balanced facet densities $f \in \mathcal{H}$ such that*

$$\frac{\ell}{2}\|f\|_{\mathcal{H}}^2 \le 1, \qquad s_m^{-1}\int xe_i^{(m)}(x)d\gamma_n(x) \longrightarrow (\mathcal{D}f)_i.$$

Proof. First we control the tails so that the normalized transfers retain their masses and first moments. We then show that they concentrate on the facets with no junction mass. Finally, we pass the sharp one-sided normal cost to the limiting densities.

Control of the tails. Normalize the signed measures by $\mu_{m,i} = e_i^{(m)}\gamma_n/s_m$. Corollary 4.4 bounds their weighted total variations and gives the uniform tail bound

$$\sup_m \sum_i \int_{|x|>R} w\, d|\mu_{m,i}| \le C\sqrt{\beta(R)} \longrightarrow 0. \tag{10}$$

The limiting measures lie on facets. On each whole pair plane define

$$\theta_{m,ij}(z) = \frac{1}{s_m\kappa}\int_0^\infty h_{ij}^{(m)}(z+tn_{ij})\phi_1(t)dt. \tag{11}$$

These nonnegative densities are bounded in $L^2(\nu_{ij})$. For the normal integral over $t < 1$ use (6); the square of its part over $t \geq 1$ is bounded by a constant times that mass. Their sum is controlled by $\int \min\{t, 1\} h \phi_1 dt$. Equation (4) proves the claim after division by s_m^2. Pass simultaneously to weak limits $\theta_{ij} \geq 0$.

For every fixed $b > 0$, (4) gives

$$s_m^{-1} \int_{t \geq b} h_{ij}^{(m)} d\gamma_n \leq C_b s_m \to 0. \tag{12}$$

For a compact continuous test function, its value at $z + tn_{ij}$ can therefore be replaced by its value at z: use uniform continuity for $t < b$, bounded normalized mass, and then let $b \downarrow 0$. Thus the ordered measures converge to $\theta_{ij}\nu_{ij}$. Their support is in F_{ij}, since any compact subset of $H_{ij} \setminus F_{ij}$ has positive distance from T_i and requires a fixed normal displacement. Junctions are plane-null sets and acquire no mass from these L^2 densities; there are no junctions when $n = 1$.

Put $f_{ij} = \theta_{ji} - \theta_{ij}$ for $i < j$ and extend antisymmetrically. Here $H_{ji} = H_{ij}$ and $n_{ji} = -n_{ij}$, so both ordered limits use the same surface measure. Then

$$\mu_{m,i} \rightharpoonup \mu_i = \sum_{j \neq i} f_{ij} \nu_{ij}|_{F_{ij}}. \tag{13}$$

The weighted tails allow passage to the limit in masses and first moments. The zero means give balance and the limiting moments are the rows of $\mathcal{D}f$.

The sharp normal cost. Fix a compact tangential cutoff $0 \leq \chi \leq 1$ and $0 < b < 1$. On the compact normal cylinder over $\operatorname{supp} \chi$, exact vanishing on the pair plane and local gradient convergence give

$$U_{\rho_m, ij}(z + tn_{ij}) \geq \rho_m(\ell - o(1))t \qquad (0 \leq t \leq b)$$

The error is relative to t, even for transfers much thinner than ρ_m. Write $m_{m,ij}^b(z) = \int_0^b h_{ij}^{(m)}(z + tn_{ij})\phi_1(t)dt$. The density bound is

$$\int_0^b t h_{ij}^{(m)}(z + tn_{ij})\phi_1(t)dt \geq \frac{(m_{m,ij}^b(z))^2}{2\kappa}.$$

Discard nonnegative loss outside the cylinders and divide by $\rho_m s_m^2 = L_{\rho_m}(h^{(m)})$ to obtain explicitly

$$1 \geq \frac{\ell - o(1)}{2} \sum_{i \neq j} \int \chi(z) \left(\frac{m_{m,ij}^b(z)}{s_m \kappa} \right)^2 d\nu_{ij}(z).$$

The truncated densities have the same weak L^2 limits as (11): their omitted normalized mass tends to zero by (12), and their L^2 norms stay bounded. Approximation of L^2 tests by bounded tests justifies this passage. Weak lower semicontinuity and exhaustion by the cutoffs give

$$1 \geq \frac{\ell}{2} \sum_{i \neq j} \|\theta_{ij}\|_{L^2(\nu_{ij})}^2 \geq \frac{\ell}{2} \|f\|_{\mathcal{H}}^2. \tag{14}$$

The last step uses $(\theta_{ji} - \theta_{ij})^2 \leq \theta_{ij}^2 + \theta_{ji}^2$. The factor is exactly $\ell/2$ because $d\nu_{ij} = \kappa \, d\gamma_{n-1}$. □

6. The comparison in every dimension

For a given A, choose the reference simplex T with moment matrix closest to B. Rotate its active space to E and write the ambient space as $E \oplus \mathbb{R}^{d-n}$, with points (x, z). As a map $E \oplus \mathbb{R}^{d-n} \to E$, the moment difference $B - B_*$ has a symmetric active block $E \to E$ and a zero transverse block. To see this, differentiate squared distance at the closest matrix along orthogonal rotations. Active rotations give all skew matrices on E as tangent vectors. Rotations mixing E and $E^\perp$ give every centered transverse block, because B_* is a times the identity on E. Orthogonality to these directions gives the two assertions.

Proposition 6.1. *For each $0 < c < 2/k$ there are $\eta, \rho_0, C > 0$ depending only on k, c such that every balanced A in $d \geq n$ with $d_k(A) < \eta$ satisfies, for $0 < \rho \leq \rho_0$ and the closest moment reference T,*

$$S_\rho(T) - S_\rho(A) \;\geq\; c\mathcal{L}_\rho(A;T) \;\geq\; C\rho\, d_k(A)^2.$$

Proof. Consider a sequence with $d_k(A) \to 0$ and $\rho \to 0$, allowing the dimension to vary. Its aligned label distance tends to zero. Indeed, choose simplices T^0 approaching A in label measure. Projection of each indicator difference onto the first Hermite degree shows

$$\|B - B_{T^0}\|_F^2 \leq \sum_i \gamma_d(A_i \triangle T_i^0), \qquad \|B_* - B_{T^0}\|_F \leq 2\|B - B_{T^0}\|_F.$$

By (1) the score regret of T^0 against B_* is $\frac{1}{2}\|B_* - B_{T^0}\|_F^2$. The same Gaussian strip bound therefore makes the label distance from T^0 to T tend to zero, uniformly in d.

We can restrict to $A \neq T$. Put $s^2 = \mathcal{L}_\rho(A;T)/\rho > 0$. By Lemma 4.3, $|U_{\rho,ij}(x)|/\rho \leq C_k w(x)$ because the score difference has zero constant term. The changed-label mass tends to zero and w is integrable in the fixed active space, so $s \to 0$. Average the transfers $h_{ij}(x,z)$ over z. These averages retain their cost, take values in $[0,1]$, and are supported in T_i. Proposition 5.1 then gives balanced facet densities f, with $\ell\|f\|_{\mathcal{H}}^2 \leq 2$. The active first moments of e/s converge to $\mathcal{D}f$, which is symmetric by the moment alignment.

Write $u_{i,p,q}$ for the component of e_i/s of Hermite degree p in x and q in z. Since T_ρ multiplies this component by ρ^{p+q}, the interaction in (8) is

$$\frac{\widetilde{Q}_\rho(e)}{\rho s^2} = \sum_{i,\, p+q \geq 1} \rho^{p+q-1}\|u_{i,p,q}\|_2^2.$$

The projection onto constants removes degree zero. The transverse first-degree component $u_{i,0,1}$ vanishes by the alignment of B_*. We bound the terms of total degree at least two by an estimate in the fixed active space.

Set $J_i(z) = s^{-1}\int w(x)|e_i(x,z)|\, d\gamma_n(x)$. Apply Corollary 4.4 in each slice and integrate in z. Since $w \geq 1$, Cauchy–Schwarz in z also bounds the changed mass:

$$\sum_i \|J_i\|_2^2 \leq C_k, \qquad \sum_i \gamma_d(A_i \triangle T_i) \leq s\sum_i \|J_i\|_2 \leq C_k s. \tag{15}$$

The L^2 norms of J_i are taken in the transverse Gaussian space.

For $r = 1/16$ the active Mehler kernel satisfies $M_r(x,y) \leq C_k w(x)w(y)$ by (3). Consequently,

$$\sum_{i,p,q} r^p \|u_{i,p,q}\|_2^2 = \sum_i \int \left\langle e_i(\cdot,z)/s, T_r(e_i(\cdot,z)/s)\right\rangle_{L^2(\gamma_n)} d\gamma_{d-n}(z) \leq C_k \sum_i \|J_i\|_2^2.$$

In $\widetilde{Q}_\rho(e)/(\rho s^2)$ the weight of a term of total degree $p+q \geq 2$ is $\rho^{p+q-1} \leq \rho r^{p+q-2} \leq (\rho/r^2) r^p$ for $\rho \leq r$. Thus these terms together are $O_k(\rho)$, uniformly in d. Only the active first moments survive in the limit. Along the extracted subsequence, Lemma 3.1 gives

$$\frac{\widetilde{Q}_\rho(e)}{\mathcal{L}_\rho(A;T)} \longrightarrow \|\mathcal{D}f\|_F^2 \leq 2(1 - 1/k).$$

If the claimed first inequality failed for every η, ρ_0, we could choose a sequence as above with interaction divided by loss greater than $2 - c$, by (8). Its limit contradicts $c < 2/k$. Choose $\rho_0 \leq r$. The second inequality follows from (15), whose constant is uniform on this range. □

Proof of Theorem 1.1. For any balanced partition, the Hermite expansion gives

$$S_\rho(A) = \frac{1}{k} + \rho W_1(A) + R_\rho(A), \qquad 0 \leq R_\rho(A) \leq (1 - 1/k)\rho^2.$$

Fix the neighborhood $d_k(A) < \eta$ from Proposition 6.1. Outside it, uniform localization in Proposition 2.1 gives a constant $\epsilon > 0$ with $F - W_1(A) \geq \epsilon$. There the Hermite estimate gives $S_\rho(T) - S_\rho(A) \geq \rho\epsilon/2$ for all sufficiently small ρ. Since $d_k(A) \leq 2$, this gives the required quadratic bound outside the neighborhood. Proposition 6.1 gives it inside. For fixed d the simplex orbit is compact in label distance: the orthogonal group is compact and convergence of rotations gives convergence of labels off the limiting pair planes. Hence zero distance is equivalent to being a simplex partition. □

7. A tail example

The first-moment deficit localizes near maximizers, but the local stability bound uses a different scale. An exchange far inside two cells explains why.

Proposition 7.1. *Fix $k \geq 2$ and $0 < \rho < 1$. There is no $c(\rho) > 0$ such that*

$$S_\rho(T) - S_\rho(A) \geq c(\rho)\rho\,[F - W_1(A)]$$

for all balanced partitions with positive first-moment deficit, even in arbitrarily small Gaussian-measure neighborhoods of T in $\mathbb{R}^{k-1}$.

Proof. Choose a unit direction v strictly inside T_i and a distinct label j. Let G reflect in H_{ij}. For large R, the ball $E_R = B(Rv, 1)$ lies in T_i, and $GE_R \subset T_j$ has the same mass p_R. Exchange the two balls. This gives a balanced partition A_R approaching T. Put $a_R = \int_{E_R} n_{ij} \cdot x\, d\gamma_n$; then $a_R \geq (Rn_{ij} \cdot v - 1)p_R$ and $a_R \to 0$. The two moment changes are $-2a_R n_{ij}$ and $2a_R n_{ij}$, so

$$F - W_1(A_R) = 4\ell a_R - 8a_R^2 \geq 2\ell a_R > 0$$

for large R. Only points in the two balls change labels; agreement can change only if at least one observation lies there. Hence $|S_\rho(T) - S_\rho(A_R)| \leq 4p_R$. Dividing by the displayed positive deficit times ρ gives a ratio tending to zero. □

For these remote changes, linear score loss grows like distance times mass, while OU score loss saturates. Normalization by the latter keeps their cost in the compactness argument, without restricting their size or location.

AI assistance and review. ChatGPT (OpenAI) assisted with mathematical arguments, literature checks, exposition, proof review and verification code.

References

[1] C. Borell, *Geometric bounds on the Ornstein–Uhlenbeck velocity process*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **70** (1985), 1–13. doi:10.1007/BF00532234.

[2] S. Heilman, *Euclidean partitions optimizing noise stability*, Electron. J. Probab. **19** (2014), 1–37. doi:10.1214/EJP.v19-3083.

[3] S. Heilman, *Three candidate plurality is stablest for correlations at most* 1/10, preprint, 2023, arXiv:2306.03312v3.

[4] S. Heilman, *Sharp hardness for MAX-3-CUT and quantum MAX-CUT*, preprint, 2026, arXiv:2608.00333v1.

[5] S. Heilman and A. Tarter, *Three candidate plurality is stablest for small correlations*, Forum Math. Sigma **9** (2021), e65. doi:10.1017/fms.2021.56.

[6] M. Isaksson and E. Mossel, *Maximally stable Gaussian partitions with discrete applications*, Israel J. Math. **189** (2012), 347–396. doi:10.1007/s11856-011-0181-7.

[7] E. Milman and J. Neeman, *The Gaussian double-bubble and multi-bubble conjectures*, Ann. of Math. (2) **195** (2022), 89–206. doi:10.4007/annals.2022.195.1.2.

[8] A. Prékopa, *On logarithmic concave measures and functions*, Acta Sci. Math. (Szeged) **34** (1973), 335–343. Journal archive.

[9] Yongxian Zhang, *Equal Gaussian cells: the sharp simplex first-moment theorem*, preprint, version 2.0, 2026. doi:10.5281/zenodo.23250730.