Entropy curvature under Gaussian smoothing: restoration and moment-matched obstructions

Scott Schubr
Luminary Oddities

Abstract

For every fixed retained signal $q \in (0,1)$ and every prescribed finite Gaussian moment matching order, we construct smooth strongly log-concave symmetric inputs approaching the Gaussian whose three-copy entropy fails concavity at input-dependent interior weights tending to a face. The construction matches the fourth moment exactly and uses a higher-order face expansion with a controlled differentiated remainder. In contrast, for $m \geq 3$ copies and each $Q < m^{-1/(m-1)}$, one bounded Gaussian-relative neighborhood has a negative definite entropy Hessian in every tangent direction at the exact barycentre for non-Gaussian inputs, simultaneously for all $0 < q \leq Q$. This bound holds at the barycentre itself. For two copies and every retained signal below one half, an explicit bounded Gaussian-relative neighborhood has strict all-weight concavity; this answers Luo's Problem 10.2 negatively on $0 < q < 1/2$. We also weaken eventual two-copy restoration to finite moments: if the first differing Gaussian moment has even order $n \geq 4$, absolute moment $2n+2$ suffices for strict all-weight concavity at small signal; one common moment bound of order $2N+4$ suffices under a positive finite Hermite discrepancy through degree N. All proofs are included, together with degree-by-degree Gaussian-boundary estimates, whole-class small-signal endpoint restoration and moment-matched exponential-tail examples, which separate the pointwise, family-uniform, barycentric and near-face regimes. For every $m \geq 3$ copies we identify the sharp barycentric retained-signal threshold q_m, with a common concave weight ball below it, radial concavity at it and positive curvature above it; at q_m the full neighborhood persists for three and four copies and fails for every $m \geq 5$. The sharp two-copy threshold is $4/7$ in every fixed dimension, the multicopy thresholds persist in every dimension, and the sharp dimension-uniform Ornstein–Uhlenbeck interpolation power is the elementary spectral power, with dimension-uniform entropy neighborhoods through signal $1/2$ for two copies; at sufficiently unequal weights a uniform curvature neighborhood holds in every dimension for each signal below one, on explicit weight intervals including their endpoints, and at the endpoints without smoothing. Finally, a universal positive signal makes the two-copy entropy concave on the full weight interval for every symmetric variance-one log-concave law, answering the existence question in Luo's Problem 10.1, and in each fixed dimension for every centrally symmetric isotropic log-concave vector, with strict curvature unless the law is Gaussian.

Keywords. Entropy concavity; Gaussian smoothing; Hermite expansion; log-concave laws; probability inequalities.
Mathematics Subject Classification (2020). 60E15; 94A17.

1 Introduction and main result

This is the eighth edition of this article. The eighth edition adds Sections 14 and 15, which prove uniform entropy concavity on the full weight interval after Gaussian smoothing for every symmetric

log-concave law and for every centrally symmetric isotropic log-concave vector in each fixed dimension, and revises Section 13 with endpoint curvature without smoothing; it also places the sections added since the sixth edition in the main text. The seventh edition added Section 13, which proved dimension-uniform entropy curvature at sufficiently unequal weights for every signal below one. The sixth edition added Sections 9–12.3, which determine the sharp barycentric retained-signal threshold for every number of copies, classify critical weight neighborhoods by copy count, extend the thresholds to every fixed dimension, and identify the sharp dimension-uniform interpolation power.

Let X, Y be independent copies of a centered variance-one random variable. For an independent standard Gaussian G, write

$$X_q = \sqrt{q}\, X + \sqrt{1-q}\, G, \qquad 0 \le q \le 1.$$

Given independent copies U_q, V_q of X_q, set

$$F_X(q,t) = h(\sqrt{t}\, U_q + \sqrt{1-t}\, V_q), \qquad 0 \le t \le 1,$$

where h is differential entropy with natural logarithms. Equivalently the output is

$$Z_{q,t} = \sqrt{qt}\, X + \sqrt{q(1-t)}\, Y + \sqrt{1-q}\, G. \tag{1}$$

The variance is one, and $F_X(q,t) = F_X(q,1-t)$.

The new results distinguish restoration at balanced weights from failure near faces, even after arbitrarily high finite Gaussian moment matching. They also replace all-moment assumptions in eventual two-copy restoration by explicit sufficient finite moments. The barycentric theorem is infinitesimal; it is compatible with the near-face obstructions.

Write

$$\varphi(x) = (2\pi)^{-1/2} e^{-x^2/2}, \qquad T_a e_n = a^n e_n, \qquad N e_n = n e_n,$$

where $e_n = \mathrm{He}_n / \sqrt{n!}$ is the normalized probabilists' Hermite basis of $L^2(\varphi\, dx)$. All unqualified norms and inner products below use this Gaussian measure.

For $m \ge 3$ independent copies $X_1, \dots, X_m$ and an independent standard Gaussian G, define on the probability simplex

$$\mathcal{F}^{(m)}_{q,g}(t) = h\left(\sum_{i=1}^{m} \sqrt{q t_i}\, X_i + \sqrt{1-q}\, G \right), \qquad X_i \sim \varphi(1+g).$$

Write $u = (1/m, \dots, 1/m)$ for its barycentre.

Theorem 1.1 (Moment-matched three-copy obstructions)**.** *Fix $q \in (0,1)$ and an integer $K \ge 2$. There exist even smooth strongly log-concave variance-one densities $f_j = \varphi(1+g_j)$ whose Gaussian moments match exactly through order $2K$, with $\|g_j\|_{C^r_b} \to 0$ for every fixed integer $r \ge 0$, but whose three-copy entropy is not concave on the simplex. More precisely, there are $\epsilon_j \downarrow 0$ for which its second derivative in direction $(1,-1,0)$ at $(\epsilon_j, \epsilon_j, 1-2\epsilon_j)$ is strictly positive. The input sequence and weights may depend on fixed q, K; no common face distance or barycentric counterexample is asserted.*

Here $\|g\|_{C^r_b}$ controls g and its derivatives through order r uniformly. The theorem imposes arbitrarily high finite Gaussian matching, not agreement of the entire moment sequence.

Theorem 1.2 (Multicopy barycentric Hessian)**.** *Fix $m \geq 3$ and $0 < Q < m^{-1/(m-1)}$. There is an explicit $\delta_{m,Q} > 0$ such that every real even measurable g satisfying*

$$\|g\|_\infty \leq \delta_{m,Q}, \qquad \int g\varphi = \int x^2 g\varphi = 0$$

has, for all $0 < q \leq Q$ and all real tangent vectors v with $\sum_i v_i = 0$,

$$\operatorname{Hess} \mathcal{F}^{(m)}_{q,g}(u)[v,v] \leq -\frac{m}{16} \|v\|^2_{\mathbb{R}^m} \|NL_q\|_2^2, \qquad L_q = mT_{\sqrt{q/m}} g. \tag{2}$$

The Hessian is negative definite on the tangent space unless $g = 0$ almost everywhere. The assertion is at the exact barycentre; it does not supply a common surrounding weight region or a sharp signal threshold.

For three copies the sufficient range is $Q < 1/\sqrt{3}$. The pointwise Hessian conclusion is compatible with failures near faces: these are different points of the simplex. For each fixed input and signal, Hessian continuity also gives an input-dependent open neighborhood of the barycentre. No uniform such neighborhood is asserted here.

1.1 Two-copy local background

The central question is whether nonconcavity can persist arbitrarily close to the Gaussian after a fixed amount of Gaussian smoothing. Luo [7, Problem 10.2] asks, for a fixed signal fraction $q \in (0, 4/7]$ and an integer $K \geq 1$, whether there exist symmetric strongly log-concave variance-one densities f_j matching the Gaussian moments through order $2K$ such that

$$\|f_j/\varphi - 1\|_{C^m_b} \longrightarrow 0 \quad \text{for every fixed } m \geq 0,$$

while $F_{X_j}(q, \cdot)$ fails concavity for every j. Here C^m_b controls the function and its derivatives through order m uniformly, and φ is the standard Gaussian density. The question is a local form of the entropy-concavity problem of Ball, Nayar and Tkocz, with its formulation and attribution taken from Luo.

We give a self-contained negative answer on $0 < q < 1/2$ as a background theorem for the restoration results below. More precisely, for every prescribed upper signal bound $Q < 1/2$ we construct an explicit positive radius in the uniform relative-density norm on which concavity holds for every mixing weight and every $0 < q \leq Q$, strictly for non-Gaussian inputs. The theorem allows arbitrary bounded even relative perturbations, including infinitely many Hermite modes. Its proof controls the nonlinear entropy terms relative to the actual quadratic Hermite energy, with a Gaussian smoothing margin that is uniform through the weight endpoints. The extensions developed here replace all-moment assumptions by sufficient finite moments and distinguish the multicopy barycentric Hessian from moment-matched near-face failure.

Theorem 1.3 (Gaussian-neighborhood concavity)**.** *For every $Q \in (0, 1/2)$ there is an explicit $\delta_Q > 0$ such that the following holds. Suppose g is real, even and measurable,*

$$\|g\|_\infty \leq \delta_Q, \qquad \int g\varphi = \int x^2 g\varphi = 0,$$

and X has density $\varphi(1+g)$. For every $0 < q \leq Q$, the second weight derivative extends continuously to $[0,1]$ and

$$\partial_t^2 F_X(q,t) \leq -\frac{1}{16} \|NL(q,t)\|_2^2, \qquad L(q,t) = T_{\sqrt{qt}} g + T_{\sqrt{q(1-t)}} g. \tag{3}$$

Thus $F_X(q, \cdot)$ is strictly concave unless $g = 0$ almost everywhere. The hypothesis allows measurable bounded relative perturbations.

An explicit sufficient radius is as follows:

$$c = \frac{1 + (2Q)^{-1/2}}{2}, \qquad r = \frac{c-1}{4}, \qquad \rho = \frac{1+2r}{c}, \tag{4}$$
$$K = \frac{\sqrt{\rho}}{(1-\sqrt{\rho})^2}, \qquad A = \frac{K}{8} + 60 + \frac{1}{8r^2}, \qquad \delta_Q = \min\left\{\frac{1}{12}, \frac{8}{K}, \frac{1}{16A}\right\}.$$

This explicit sufficient radius depends on Q and decreases to zero as $Q \uparrow 1/2$.

Corollary 1.4 (Answer to the near-Gaussian counterexample question)**.** *Fix $q \in (0, 1/2)$. There is no sequence of symmetric variance-one densities f_j with $\|f_j/\varphi - 1\|_\infty \to 0$ for which every $F_{X_j}(q, \cdot)$ fails concavity. The same statement holds after imposing strong log-concavity, any prescribed finite Gaussian moment matching, or convergence of all fixed-order relative-density derivatives.*

1.2 Relation to prior work and proof mechanism

Luo [7] constructs asymmetric counterexamples for every positive signal fraction and symmetric strongly log-concave counterexamples for $q > 4/7$, with arbitrarily close Gaussian-relative approximation and prescribed finite Gaussian moment matching. Theorem 1.3 and Corollary 1.4 identify a positive relative neighborhood in the low-signal symmetric regime. The additional regularity and moment-matching requirements in Problem 10.2 are compatible with the theorem: uniform relative-density proximity alone places every sufficiently late member of such a sequence within its concavity neighborhood.

An external unrefereed candidate revised on November 11, 2026 [5] claims a uniform bounded relative-density neighborhood for the entire two-copy range $0 < q \le 4/7$, using a critical Gaussian interpolation inequality. It credits the conditional Hermite and entropy-energy scheme used in the below-half proof here. We retain that external attribution and its unrefereed status; none of its critical smoothing theorem is a premise of the proofs below. Its Section 7 constructs multicopy failures near faces, while Section 9.3 asks about barycentric and interior survival. Our barycentric theorem concerns the exact Hessian with an explicit sufficient signal range. Theorem 1.1 extends the face mechanism to arbitrary finite Gaussian moment matching, including exact fourth matching, where that candidate's leading fourth-cumulant term vanishes and its Sections 7.4(9) and 9.3 treat the question separately. The barycentric and face results are thus complementary: one fixes the weight at the centre, the other lets it approach a face.

Gaussian scale mixtures have entropy concavity [4]. Abu-Ajamieh [2] derives the exact symmetric endpoint-curvature identity, establishes endpoint estimates and a neighborhood for an auxiliary endpoint inequality, and proves an all-weight theorem for a specified fourth-Hermite perturbation family. Our local theorem treats a bounded infinite-dimensional relative-density neighborhood. The proof combines classical Gaussian/Hermite coordinates [1], conditional-expectation contraction, a complex-weight smoothing margin, and number-operator estimates to absorb the full nonlinear curvature error. Gaussian-channel identities have further classical antecedents [6].

1.3 Complementary restoration results

The Gaussian-relative theorem describes one neighborhood topology. The results below describe how all-weight concavity is restored under a common moment envelope or for a fixed input, and give a whole-class endpoint estimate. Together with exact moment-matched tail examples, they explain the distinct roles of proximity, symmetry and retained signal.

A complementary finite-Hermite criterion permits a common threshold under one sufficiently high moment bound, even when the first non-Gaussian moment varies across the input family.

Theorem 1.5 (Finite Hermite discrepancy)**.** *Fix an even $N \geq 4$, $B < \infty$ and $\varepsilon > 0$. Let $\mathcal{C}$ be a family of symmetric variance-one laws satisfying*

$$\sup_{X \in \mathcal{C}} \mathbb{E}|X|^{2N+4} \leq B.$$

Put $a_n(X) = \mathbb{E}e_n(X)$ and suppose that

$$\sum_{\substack{4 \leq n \leq N \\ n \ even}} a_n(X)^2 \geq \varepsilon^2 \qquad (X \in \mathcal{C}). \tag{5}$$

There is $q_{B,N,\varepsilon} > 0$ such that every $X \in \mathcal{C}$ has strict all-weight entropy concavity for every $0 < q \leq q_{B,N,\varepsilon}$, including continuous second derivatives at both endpoints. The family may include discrete laws and laws with some higher moments infinite.

The finite-jet estimate also quantifies the Gaussian boundary of a family with nonnegative curvature.

Corollary 1.6 (Quantitative Gaussian boundary)**.** *Let $\mathcal{C}$ be a family of symmetric variance-one laws with a common absolute-moment envelope: $\sup_{X \in \mathcal{C}} \mathbb{E}|X|^j \leq B_j < \infty$ for every integer $j \geq 1$. For every fixed even $N \geq 4$, there are $K_N < \infty$ and $q_N > 0$ such that, for every $X \in \mathcal{C}$ and $0 < q \leq q_N$,*

$$(\exists t \in [0,1] : F_{X,tt}(q,t) \geq 0) \quad \Longrightarrow \quad \sum_{\substack{4 \leq n \leq N \\ n \text{ even}}} a_n(X)^2 q^n \leq K_N q^{N+1}. \tag{6}$$

Consequently, for each fixed polynomial P and every $A > 0$, there are $K_{P,A} < \infty$ and $q_{P,A} > 0$ such that the same antecedent, with $0 < q \leq q_{P,A}$, implies

$$|\mathbb{E}P(X) - \mathbb{E}P(G)| \leq K_{P,A} q^A.$$

The constants are uniform over $\mathcal{C}$ for each fixed P, A. In particular, along any sequence $X_j \in \mathcal{C}$, $q_j \downarrow 0$, for which $F_{X_j}(q_j, \cdot)$ fails concavity, every fixed polynomial moment discrepancy is $O(q_j^A)$ for every fixed $A > 0$.

The proof in Section 5 fixes the jet order before choosing the signal interval. Thus the conclusion gives a degree-by-degree necessary condition on a changing family. For symmetric variance-one log-concave laws, their common moment envelope also gives Gaussian-only limiting behavior (Corollary 5.1).

Theorem 1.7 (Finite-moment eventual concavity)**.** *Let X be symmetric with variance one. Suppose that for some even $n = 2k \geq 4$, its moments below order n equal the Gaussian moments, $\kappa_n = \mathbb{E}X^n - \mathbb{E}G^n \neq 0$, and $\mathbb{E}|X|^{2n+2} < \infty$. There is $q_X > 0$ such that $\partial_t^2 F_X(q,t) < 0$ for every $0 < q \leq q_X$ and every $t \in [0,1]$, with continuous second derivatives at both endpoints. Moreover*

$$F_X(q,t) = h(\varphi) - \frac{\kappa_n^2}{2n!} q^n \left[t^k + (1-t)^k\right]^2 + o(q^n) \tag{7}$$

in $C^2([0,1])$ as $q \downarrow 0$. In particular every fixed symmetric non-Gaussian variance-one law with all absolute moments finite has such an eventual threshold.

This includes every fixed symmetric variance-one log-concave input and discrete symmetric inputs with all moments finite, but also genuine polynomial-tail laws. The threshold is attached to the law. The stated finite moment orders are sufficient; no minimal-moment claim is made.

The three all-weight regimes have distinct useful inputs: an explicit relative-density neighborhood, a fixed non-Gaussian law, or a positive finite Hermite discrepancy under a common moment envelope. Corollary 1.6 quantifies the changing near-Gaussian boundary. Appendix B supplies exactly moment-matched exponential-tail examples distinguishing these topologies.

Theorem 1.8 (Whole-class endpoint restoration)**.** *Let X have a symmetric log-concave density and variance one. If $0 < q \le 1/20$ and p is the density of X_q, then*

$$\partial_t^2 F_X(q,0) = \partial_t^2 F_X(q,1) \le -\frac{1}{128}\mathbb{E}\left[\left(1 + (\log p)''(X_q)\right)^2\right]. \tag{8}$$

The endpoint derivatives are continuous one-sided extensions. The bound is strict unless X is Gaussian.

This last theorem uses the endpoint identity from [2], supplemented by uniform small-signal estimates. The conclusion concerns the two endpoint derivatives.

1.4 Organization of the proofs

Section 2 proves the below-half Gaussian-neighborhood theorem directly. Sections 3–5 develop finite-moment Gaussian-gain regularity, fixed-input eventual concavity, and family-uniform finite-Hermite and Gaussian-boundary results. Section 6 proves the multicopy barycentric Hessian bound, including the active-subset interpolation and explicit radius. Section 7 proves the higher-moment face obstruction, its differentiated remainder and the bounded moment-corrected inputs. Appendix A proves the uniform endpoint estimate, using the credited endpoint identity. Appendix B constructs the exactly moment-matched exponential tails that distinguish moment proximity from bounded Gaussian-relative proximity.

2 Proof of Gaussian-neighborhood concavity

We prove Theorem 1.3. Fix $Q < 1/2$ and the constants (4), and put $\delta = \|g\|_\infty \le \delta_Q$. The conditions on g imply that its Hermite coefficients a_n vanish except at even orders $n \ge 4$. The output density from (1) is $\varphi(1+R)$, where

$$R = L + M, \qquad L = T_{\sqrt{qt}}g + T_{\sqrt{q(1-t)}}g, \tag{9}$$

and

$$M(x) = \mathbb{E}[g(G_1)g(G_2) \mid \sqrt{qt}G_1 + \sqrt{q(1-t)}G_2 + \sqrt{1-q}G_3 = x].$$

The G_i here are independent standard Gaussians. Hence $\|R\|_\infty \le 2\delta + \delta^2 \le 3\delta =: \eta \le 1/4$ on the real weight interval. The output has mass one and variance one, so

$$F_{tt} = -\langle R'', \log(1+R)\rangle - \int \frac{(R')^2}{1+R}\varphi. \tag{10}$$

Primes in this section denote weight derivatives unless otherwise specified. We justify this differentiation below after deriving the required Hilbert-space regularity.

2.1 Linear energy

For $n = 2k \geq 4$ the multiplier of a_n in L is

$$l_n = q^k S_k(t), \qquad S_k(t) = t^k + (1-t)^k.$$

Log-convexity of power sums and $S_1/S_0 = 1/2$ show $S_k \geq S_{k-1}/2$ and $S_k \geq S_{k-2}/4$; also $S_{k-2} \geq S_k$. Thus

$$|l_n'| \leq n l_n, \qquad \frac{n^2}{8} l_n \leq l_n'' \leq n^2 l_n.$$

With $D = \|NL\|_2$, we obtain

$$\langle L'', L\rangle \geq D^2/8, \qquad \left\|N^{-1}L''\right\|_2 \leq D, \qquad \|L\|_2 \leq D/4. \tag{11}$$

The inverse is restricted to positive Hermite degrees. If $g \neq 0$, then $D > 0$ at every positive q and every weight.

2.2 Bilinear term and a smoothing margin

Gaussian convolution in Hermite coordinates [1] gives

$$M_n(q,z) = q^{n/2} \sum_{j+k=n} a_j a_k \binom{n}{j}^{1/2} z^{j/2}(1-z)^{k/2}. \tag{12}$$

Only even $j, k \geq 4$ occur. Let g_{abs} be the $L^2(\varphi)$ function with coefficients $|a_n|$; it need not be bounded, but $\|g_{\mathrm{abs}}\|_2 \leq \delta$.

Fix $t \in [0,1]$ and $|z - t| \leq r$. Set $a = |z|$, $b = |1-z|$, $P = q(a+b)$, $w = a/(a+b)$ and $\bar{P} = cq$. Our choice of constants gives

$$\begin{gathered} P \leq q(1+2r) < \bar{P} < 1, \qquad P/\bar{P} \leq \rho < 1, \\ \bar{P}^2 = c^2 q^2 \leq c^2 Q q < q/2. \end{gathered} \tag{13}$$

Absolute coefficient majorization in (12) reduces the norm to the real bilinear expression $B(P, w; g_{\mathrm{abs}}, g_{\mathrm{abs}})$. For any $h \in L^2(\varphi)$,

$$\|B(P, w; h, h)\|_2 \leq \left\|T_{\sqrt{P}} h\right\|_2^2. \tag{14}$$

Indeed smooth two independent inputs separately and condition their product on their weighted Gaussian sum; L^2 contraction and independence give (14). For two inputs, the same conditioning argument bounds the bilinear map by the product of their smoothed L^2 norms. Applying this bound to differences of finite Hermite truncations gives L^2 convergence and extends the identity and inequality to L^2 inputs.

Spectral Cauchy–Schwarz gives

$$\left\|T_{\sqrt{\bar{P}}} g_{\mathrm{abs}}\right\|_2^2 = \sum a_n^2 \bar{P}^n \leq \delta \left\|T_{\bar{P}} g_{\mathrm{abs}}\right\|_2.$$

By (13) and $S_k(t) \geq 2^{1-k}$,

$$\|T_{\bar{P}} g_{\mathrm{abs}}\|_2 \leq \left\|T_{\sqrt{q/2}} g\right\|_2 \leq \|L\|_2 / 2.$$

Consequently $\|M(q,z)\|_2 \le \delta \|L\|_2/2$ throughout this disk. For NM, split off the output multiplier $(P/\bar{P})^{n/2}$; its number-operator cost is at most

$$\sup_{n\ge 0} n\rho^{n/2} \le \sum_{n\ge 1} n(\sqrt{\rho})^n = K.$$

It follows that $\|NM(q,z)\|_2 \le K\delta \|L\|_2/2$. Strictness in (13) leaves a slightly larger disk. The same majorants, with the exponentially decaying output multiplier, show normal convergence of the coefficient series and all fixed number-operator derivatives there. Hence M is Hilbert-valued analytic. Cauchy's estimate on radius r and (11) yield

$$\|M\|_2 \le \delta D/8, \qquad \|NM\|_2 \le K\delta D/8, \qquad \|M''\|_2 \le \delta D/(4r^2). \tag{15}$$

The linear term is analytic too, since $q|z|$ and $q|1-z|$ are below $\bar{P} < 1$. All estimates are uniform in $t \in [0,1]$ and $0 < q \le Q$.

2.3 A bounded-function estimate

For smooth bounded real u with $Nu \in L^2(\varphi)$,

$$\left\|(u_x)^2\right\|_2 \le 3\|u\|_\infty \|Nu\|_2 . \tag{16}$$

To prove it, put $B = \int u_x^4 \varphi$. Gaussian integration by parts gives

$$\int Nu\, u(u_x)^2\varphi = B + 2\int u(u_x)^2 u_{xx}\varphi.$$

Cauchy–Schwarz and the Gaussian Bochner identity $\|u_{xx}\|_2 \le \|Nu\|_2$ imply $B \le 3\|u\|_\infty \sqrt{B}\,\|Nu\|_2$, proving the assertion. Smooth cutoffs and Gaussian Sobolev approximation justify the operator-domain version; here positive smoothing also bounds the relevant spatial derivatives.

Set $\psi(v) = \log(1+v) - v$. The chain rule gives

$$N\psi(R) = \psi'(R)NR - \psi''(R)(R_x)^2.$$

Using (16) and $\|R\|_\infty \le \eta \le 1/4$,

$$\|N\psi(R)\|_2 \le \left[\frac{\eta}{1-\eta} + \frac{3\eta}{(1-\eta)^2}\right]\|NR\|_2 \le 30\delta(1 + K\delta/8)D. \tag{17}$$

2.4 Curvature sign and regularity

The density ratio lies in $[1-\eta, 1+\eta]$. Hilbert-valued C^2 regularity of R and the bounded second derivative of $u\log u$ on this interval justify (10): scalar Taylor's formula and L^2 convergence of difference quotients give the differentiated integrals. The resulting formula is continuous to the endpoints. Spatial regularity from smoothing and the preceding number-operator estimates justify self-adjointness and the chain rule used in (17).

Drop the last, nonpositive term of (10) and decompose $\log(1+R) = L + M + \psi(R)$. The error terms satisfy

$$\begin{aligned} |\langle L'', M\rangle| &\le K\delta D^2/8, \\ |\langle L'', \psi(R)\rangle| &\le 30\delta(1 + K\delta/8)D^2 \le 60\delta D^2, \\ |\langle M'', \log(1+R)\rangle| &\le \frac{\delta D^2}{4r^2}\frac{1/4 + \delta/8}{1 - 3\delta} \le \frac{\delta D^2}{8r^2}. \end{aligned}$$

Here $\delta \le 8/K$ in the second estimate, and $\delta \le 1/12$ gives $(1/4+\delta/8)/(1-3\delta) \le 25/72 < 1/2$ in the third. Therefore

$$F_{tt} \le -D^2/8 + A\delta D^2 \le -D^2/16.$$

This proves Theorem 1.3. Corollary 1.4 follows by taking $Q = q$ and observing that every sufficiently large member of the proposed sequence satisfies its relative-density hypothesis.

3 Finite-moment Gaussian-gain regularity

Lemma 3.1. *Let $k \ge 1$ be an integer and suppose $\mathbb{E}|X|^{k+2} < \infty$. For independent copies X, Y and an independent standard Gaussian G, put, for $H = (h_1, h_2)$ with $|H| < 1$,*

$$\sigma(H) = \sqrt{1 - h_1^2 - h_2^2}, \qquad Z_H = h_1 X + h_2 Y + \sigma(H) G, \qquad K(H) = h(Z_H).$$

Then K is C^k. On a fixed compact gain disk its derivatives through order k have common bounds under a common $(k+2)$nd absolute-moment bound and a common positive bounded-input mass. For centered variance-one laws the latter mass is uniform. If X is symmetric, K is even in each coordinate. When $k = 2s$, the function $J(a,b) = K(\sqrt{a}, \sqrt{b})$ is C^s up to the axes of a small positive quadrant, with the corresponding common derivative bounds.

Proof. Work on a fixed closed disk where σ stays positive. The density is

$$p_H(z) = \mathbb{E}\varphi_{\sigma(H)}(z - h_1 X - h_2 Y).$$

For a gain multiindex α, $j = |\alpha| \le k$, kernel differentiation gives

$$\frac{\partial_H^\alpha p_H(z)}{p_H(z)} = \mathbb{E}[P_{\alpha,H}(X, Y, G) \mid Z_H = z], \tag{18}$$

where the coefficients are smooth and bounded on the disk, the input degree in (X, Y) is at most j, and the Gaussian-residual degree is at most $2j$. To verify these separate degree bounds, let $u = (z - h_1 X - h_2 Y)/\sigma$. One kernel derivative divided by the kernel is

$$X_i u/\sigma + (\sigma_i/\sigma)(u^2 - 1), \qquad X_1 = X, \quad X_2 = Y.$$

A derivative of a previously formed polynomial also uses $\partial_{h_i} u = -X_i/\sigma - (\sigma_i/\sigma)u$. Each step adds at most one input degree and two residual degrees. Replacing the residual by the actual independent G gives (18). On compact spatial intervals differentiation under the expectation is dominated by at most j input factors, since Gaussian-polynomial kernel derivatives are bounded.

A positive-mass bounded rectangle in (X, Y) gives constants, uniform on the gain disk, such that

$$p_H(z) \ge c e^{-C(1+z^2)}, \qquad p_H(z) \le C', \qquad |\log p_H(z)| \le C''(1+z^2).$$

For variance-one centered inputs, $\Pr(|X| \le 2) \ge 3/4$, so these constants can be common to the family.

Every formal entropy derivative of order j is a sum of a term $\partial_H^\alpha p_H(1 + \log p_H)$ and terms $p_H \prod_i r_{\beta_i,H}$, where $r_\beta = (\partial_H^\beta p_H)/p_H$, $b_i = |\beta_i| \ge 1$ and $\sum_i b_i = j$. Conditional Jensen and weighted Young give

$$\prod_i |r_{\beta_i,H}| \le \sum_i \frac{b_i}{j} |r_{\beta_i,H}|^{j/b_i} \le \mathbb{E}\left[\sum_i \frac{b_i}{j} |P_{\beta_i,H}|^{j/b_i} \,\middle|\, Z_H = z \right].$$

Their joint majorant has input degree at most j and Gaussian degree at most $2j$. The logarithmic term has a majorant

$$C(1+|X|+|Y|)^{j+2}(1+|G|)^{2j+2}.$$

This is integrable by independence and the $(k+2)$nd input moment. Uniformly on the gain disk, $|Z_H| > R$ implies $C_1(1+|X|+|Y|+|G|) > R$. Dominated convergence for this fixed joint law therefore removes all spatial tails uniformly in H. On compact spatial sets the density derivatives are continuous and denominators are positive.

Insert smooth compact spatial cutoffs into the entropy integral. The truncated integrals are C^k; their derivatives through order k converge uniformly on the gain disk by the preceding tail bounds. The order-zero tails use the same quadratic logarithm and second moments. Uniform derivative convergence along coordinate segments proves actual C^k regularity and identifies the derivatives with their integrals. Full-expectation bounds without a tail indicator give the claimed family-uniform derivative bounds. No uniform integrability of a family's highest moments is asserted.

Symmetry gives evenness in each gain. If E is a smooth even function and $a = h^2$, the finite-order formula

$$\frac{d^j}{da^j}E(\sqrt{a}) = \frac{1}{2^{2j-1}(j-1)!}\int_0^1 (1-u^2)^{j-1}E^{(2j)}(uh)\,du \tag{19}$$

holds whenever E is C^{2j}. It follows for $j = 1$ by the fundamental theorem of calculus and then by induction and integration by parts. At $a = 0$ it equals $j!E^{(2j)}(0)/(2j)!$. Applying it in both coordinates uses total gain order at most twice the squared-gain order, and proves the C^s conclusion, including mixed derivatives on the axes. Taylor formulas follow along segments in the positive quadrant. □

4 Eventual all-weight concavity for each input

We prove the finite-moment assertion first. Its specified first differing even moment is $n = 2k \geq 4$; the cumulants of orders $3, \ldots, n-1$ vanish and $\kappa_n = \mathbb{E}X^n - \mathbb{E}G^n \neq 0$. Lemma 3.1 with gain order $2n$ gives $K \in C^{2n}$ and squared-gain regularity $J \in C^n$, using only moment $2n+2$.

Use the gain model from Lemma 3.1. For its characteristic function,

$$\mathbb{E}e^{i\xi Z_H} = e^{-\xi^2/2}[e^{h_1^2\xi^2/2}\chi_X(h_1\xi)][e^{h_2^2\xi^2/2}\chi_X(h_2\xi)].$$

Finite moment differentiation shows that the density gain jet has no nonconstant term below total degree n, and its degree-n term is

$$p_{[n]}(H,z) = \frac{\kappa_n}{n!}(h_1^n + h_2^n)\varphi(z)\,\mathrm{He}_n(z). \tag{20}$$

This identifies a finite Taylor jet, not a convergent cumulant series. All density derivatives used in this finite jet are Gaussian times polynomials, so Fourier-transform uniqueness verifies the density identification directly.

Write $D(H) = D(p_H\|\varphi)$. By Lemma 3.1, its derivatives through total gain order $2n$ can be computed under the integral. At $p = \varphi$ the integrand jet is

$$p\log(p/\varphi) = (p-\varphi) + \frac{(p-\varphi)^2}{2\varphi} + \text{terms of at least third order in } p-\varphi.$$

This is used only for finite gain derivatives at zero. The integrated linear term vanishes by mass conservation. Three nonzero density derivatives have total gain degree at least $3n$, so through

degree $2n$ the only nonlinear contribution is the square of (20). Since $\int \mathrm{He}_n^2\,\varphi = n!$,

$$D(H) = \frac{\kappa_n^2}{2n!}(h_1^n + h_2^n)^2 + R_0(H). \tag{21}$$

Here R_0 has zero gain jet through total degree $2n$. The output has mean zero and variance one, hence $K(H) = h(\varphi) - D(H)$.

In squared gains $a = h_1^2, b = h_2^2$, the even-square formula (19) identifies Taylor coefficients at the origin and gives

$$J(a,b) = h(\varphi) - \frac{\kappa_n^2}{2n!}(a^k + b^k)^2 + R(a,b).$$

The Taylor polynomial of R through total degree n vanishes. Taylor's theorem for its derivatives yields

$$\partial^\alpha R(a,b) = o\big((a+b)^{n-|\alpha|}\big), \qquad |\alpha| \le 2,$$

uniformly over positive-quadrant directions, including the axes. Substituting $a = qt, b = q(1-t)$ proves (7) in $C^2([0,1])$: two weight derivatives introduce $q^2(R_{aa} - 2R_{ab} + R_{bb})$.

For $S_k(t) = t^k + (1-t)^k$ and $k \ge 2$,

$$S_k \ge 2^{1-k}, \quad S_k'' \ge k(k-1)2^{3-k}, \quad (S_k^2)'' \ge k(k-1)2^{5-2k} =: c_k > 0.$$

For $k = 2$ the second bound uses $S_0 = 2$. The uniform little-oh remainder permits choosing $q_X > 0$ so that

$$\partial_t^2 F_X(q,t) \le -\frac{\kappa_n^2 c_k}{4n!} q^n < 0$$

for all $0 < q \le q_X$ and all weights. This proves Theorem 1.7. For the all-moment assertion, Gaussian moments are determinate: $(2j-1)!! \le (2j)^j$ makes the Carleman sum diverge. A non-Gaussian symmetric all-moment law therefore has a first differing even moment, so the finite-moment theorem applies. Log-concave laws have exponential tails and hence all absolute moments.

For a genuinely finite-moment example, consider

$$f(x) = \frac{256}{189\pi}(1 + x^2/9)^{-6}.$$

The beta integral gives mass one, variance one and fourth moment $27/7$, so $\kappa_4 = 6/7$. Its tail is proportional to $|x|^{-12}$, giving finite tenth absolute moment and infinite eleventh absolute moment. Thus the theorem with $n = 4$ applies, although the all-moment hypothesis does not. This standardized Student law is not globally log-concave and is not a whole log-concave-class example.

5 Uniform finite-moment restoration and the Gaussian boundary

We prove Theorem 1.5 by a relative finite entropy-jet estimate that allows changing lower cumulants, then extract its quantitative Gaussian boundary.

Proof. Uniform finite derivatives. On the common gain disk $|H| \le 1/2$, Chebyshev's inequality gives $\Pr(|X| \le 2) \ge 3/4$. Thus the positive kernel lower bound in Lemma 3.1 is uniform, as are the density upper bound and quadratic logarithm bound. The single moment bound $2N + 4$ in Theorem 1.5 bounds all lower moments by Lyapunov. Lemma 3.1 at gain order $2N + 2$ gives

uniform entropy derivative bounds through that order. Formula (19) transfers them to squared-gain derivatives through order $N+1$ on a common small quadrant. If $P_{X,N}$ denotes the degree-N Taylor polynomial of J_X, then

$$|\partial^\alpha (J_X - P_{X,N})(a,b)| \le C_N (a+b)^{N+1-|\alpha|}, \qquad |\alpha| \le 2. \tag{22}$$

Here and below constants depend only on B and N. The integral Taylor formula applied to the derivatives uses uniform bounds through degree $N+1$, not a common modulus of continuity of the highest derivative across the family. Substitution $a = qt, b = q(1-t)$ bounds the remainder and each of its first two weight derivatives by a common constant times q^{N+1}.

The finite entropy jet. Only a formal finite jet at zero is used. By Gaussian-kernel differentiation, the relative-density jet is $1 + L + M$ with positive-degree Hermite coefficients

$$L_n = a_n(X)(a^{n/2} + b^{n/2}),$$
$$M_n = \sum_{\substack{j+k=n \\ j,k \ge 4 \text{ even}}} a_j(X) a_k(X) \binom{n}{j}^{1/2} a^{j/2} b^{k/2}.$$

Each symbol $a_n(X)$ carries squared-gain degree $n/2$. Mass conservation removes the linear relative-entropy term. Hermite orthogonality makes its quadratic contribution from L alone, through total squared-gain degree N, exactly

$$\frac{1}{2} \sum_{\substack{4 \le n \le N \\ n \text{ even}}} a_n(X)^2 (a^{n/2} + b^{n/2})^2. \tag{23}$$

Every other nonzero monomial in this finite relative-entropy jet has at least three input factors, because LM has three, M^2 has four, and higher entropy powers have at least three. Such a term is a numerical constant times

$$\left(\prod_{\ell=1}^{m} a_{n_\ell}(X) \right) a^r b^s, \quad m \ge 3, \quad n_\ell \ge 4 \text{ even}, \quad r + s = \frac{1}{2} \sum_{\ell=1}^{m} n_\ell \le N. \tag{24}$$

Index conservation follows directly from the density jet and multiplication; integration of Gaussian polynomials only changes the numerical constants. There are finitely many such monomials, and r, s are nonnegative integers. These statements compute finite derivatives justified above.

Relative control of nonlinear terms. Set

$$E_{X,N}(q) = \sum_{\substack{4 \le n \le N \\ n \text{ even}}} a_n(X)^2 q^n.$$

Order the indices in (24) as $n_1 \le n_2 \le \cdots \le n_m$. Necessarily $n_2 \le N$: otherwise already $n_2 + n_3 > 2N$, contrary to index conservation. Thus the two smallest factors belong to $E_{X,N}$. For $0 < q \le 1$,

$$|a_{n_1} a_{n_2}| q^{(n_1+n_2)/2} \le \frac{1}{2} \left(a_{n_1}^2 q^{n_1} + a_{n_2}^2 q^{n_2} \right).$$

The remaining coefficients have common bounds. Their factor is q^β with $\beta \ge 2$, so it is bounded above by q^2 for $0 < q \le 1$. The second derivative of $t^r (1-t)^s$ is bounded on $[0,1]$ by a monomial-dependent constant. Summing the finite monomial list therefore bounds the absolute nonlinear contribution to weight curvature by $C_N q^2 E_{X,N}(q)$, uniformly through the endpoints.

Coercivity. For $n = 2k$, Section 4 gives $(S_k^2)'' \geq k(k-1)2^{5-2k} > 0$. Let

$$c_N = \frac{1}{2} \min_{2 \leq k \leq N/2} k(k-1)2^{5-2k} > 0.$$

Since entropy is Gaussian entropy minus relative entropy, equations (22)–(23) imply

$$F_{tt}(q,t) \leq -c_N E_{X,N}(q) + C_N q^2 E_{X,N}(q) + C_N q^{N+1}.$$

By (5), $E_{X,N}(q) \geq \varepsilon^2 q^N$. Choose a common positive threshold below the gain-domain bound and below one so that $C_N q^2 \leq c_N/4$ and $C_N q \leq c_N \varepsilon^2/4$ at and below it. Then

$$F_{tt}(q,t) \leq -\frac{c_N}{2} E_{X,N}(q) < 0,$$

uniformly over the entire family and all weights, proving the theorem. □

Proof of Corollary 1.6. Fix an even $N \geq 4$. The preceding proof establishes the uniform curvature inequality

$$F_{X,tt}(q,t) \leq -c_N E_{X,N}(q) + C_N q^2 E_{X,N}(q) + C_N q^{N+1}$$

on a common small-signal interval under the moment envelope, with $c_N > 0$. Choose that interval smaller so that $C_N q^2 \leq c_N/2$. At a point of nonnegative curvature,

$$E_{X,N}(q) \leq \frac{2C_N}{c_N} q^{N+1}.$$

Every summand is nonnegative, so for even $4 \leq n \leq N$,

$$|a_n(X)| \leq \left(\frac{2C_N}{c_N}\right)^{1/2} q^{(N+1-n)/2}. \tag{25}$$

For each fixed even $n \geq 4$ and $A > 0$, choose one even $N \geq \max\{4, n, n + 2A - 1\}$ and take $q \leq 1$ as well. Then (25) is $O(q^A)$ uniformly over $\mathcal{C}$. The moments of degrees zero, one and two already have their Gaussian values, and symmetry fixes every odd moment. Each fixed polynomial is a finite linear combination of Hermite polynomials. Taking the smallest of the finitely many signal thresholds and summing their constants proves the polynomial assertion. Finally Gaussian-gain regularity makes the weight function C^2 on the closed interval in this regime. Its failure of concavity supplies an interior point of positive second derivative. Applying the same estimate at that point proves the sequence assertion. □

Corollary 5.1 (Gaussian-only limiting obstruction)**.** *Let $\mathcal{L}$ be the class of symmetric variance-one log-concave laws on $\mathbb{R}$, and let $\mathcal{U}$ be any weakly open neighborhood of the standard Gaussian law γ. There is $q_{\mathcal{U}} > 0$ for which every $X \in \mathcal{L} \setminus \mathcal{U}$ has strict all-weight entropy concavity whenever $0 < q \leq q_{\mathcal{U}}$. Consequently, if $X_j \in \mathcal{L}$, $q_j \downarrow 0$, and $F_{X_j}(q_j, \cdot)$ fails concavity for every j, then X_j converges weakly to γ and in every fixed absolute moment.*

Proof. The log-concave moment estimate $\|X\|_p \leq p/2$ for $p \geq 2$ [9] supplies a common moment envelope. These laws are tight; uniform higher moments preserve variance and every fixed moment under weak convergence. Log-concavity and symmetry are closed under weak limits [10], and preserved variance excludes a degenerate limit. Thus $\mathcal{L}$ is weakly compact and all fixed Hermite moments are continuous on it.

The compact set $\mathcal{L} \setminus \mathcal{U}$ excludes γ. For each of its members, some even Hermite moment of order at least four is nonzero, by Gaussian moment determinacy. The positive loci of the

increasing continuous functions $\sum_{4\le n\le N,\, n \text{ even}} a_n(X)^2$ cover this compact set. A finite subcover gives a common even N, and the positive minimum of the resulting function gives (5). Theorem 1.5 supplies the uniform threshold. If a proposed bad sequence did not converge weakly to γ, a subsequence would stay outside one weak Gaussian neighborhood and contradict this threshold. Uniform higher moments then give convergence of every fixed absolute moment. □

Fixing a first differing moment order and a positive cumulant gap is one special case of Theorem 1.5. The finite-jet criterion also permits lower moments to vary and lower cumulants to vanish. Corollary 1.6 supplies the quantitative necessary condition as the discrepancy shrinks, degree by degree.

6 The multicopy Hessian at the barycentre

We prove Theorem 1.2 with all active subset sizes retained in the smoothing estimate. The constants in this section are distinct from those of the two-copy local proof. Put

$$c_0 = \min_{2\le j\le m}\left[\frac{m^{j-1}}{j^j Q^{j-1}}\right]^{1/j}, \quad c = \frac{1+c_0}{2}, \quad r = \frac{(c-1)\sqrt{2}}{4m},$$
$$\rho = \frac{1+mr/\sqrt{2}}{c}, \quad K = \frac{\sqrt{\rho}}{(1-\sqrt{\rho})^2}, \quad S = \frac{2^m-m-1}{m},$$
$$a = 2^m-1, \quad A = \frac{mKS}{16}+\frac{7ma}{2}+\frac{S}{3r^2}, \quad \delta_{m,Q} = \min\left\{1, \frac{1}{4a}, \frac{4}{KS}, \frac{1}{S}, \frac{m}{16A}\right\}. \tag{26}$$

These are finite positive constants. Indeed $x^{x/(x-1)}$ increases strictly for $x>1$, since its logarithmic derivative is $(x-1-\log x)/(x-1)^2>0$. Hence the smallest of $m/j^{j/(j-1)}$, $2\le j\le m$, is $m^{-1/(m-1)}$. Our hypothesis implies $c_0>1$, and simultaneously

$$c^j Q^{j-1} j^j / m^{j-1} < 1 \qquad (2\le j\le m). \tag{27}$$

The minimum over j is essential; controlling only the last subset size after choosing a large c would not imply the other inequalities.

6.1 Linear energy and the conditional subset expansion

Normalize v to Euclidean norm one, with $\sum_i v_i = 0$, and let $t_i(s) = 1/m + sv_i$. Write $\delta = \|g\|_\infty$. For real s sufficiently near zero these are positive simplex weights. Conditioning independent Gaussian input factors gives output density $\varphi(1+R_s)$, where

$$R_s = L_s + M_s, \qquad L_s = \sum_i T_{\sqrt{qt_i(s)}} g,$$

and M_s is the sum over all subsets $H\subseteq[m]$ with $|H|\ge 2$ of the conditional products of the g factors in H. The Hermite coefficient of one such product, at total degree $n=\sum_{i\in H} n_i$, is

$$q^{n/2}\sum_{\sum n_i = n}\left(\frac{n!}{\prod_{i\in H} n_i!}\right)^{1/2}\prod_{i\in H} a_{n_i} t_i^{n_i/2}. \tag{28}$$

Every active n_i is even and at least four. The formula follows by conditioning products of normalized Hermites on their Gaussian sum; it also follows by multiplying their Gaussian generating functions. On real simplex weights,

$$\|R_s\|_\infty \le (1+\delta)^m - 1 \le a\delta =: \eta \le 1/4. \tag{29}$$

The second inequality uses $\delta^j \le \delta$ for $\delta \le 1$. The output retains mass one and variance one.

At $s = 0$, put $L = L_0 = mT_{\sqrt{q/m}}g$ and $D = \|NL\|_2$. For $n = 2k \ge 4$, the multiplier of a_n in L_s is $l_n(s) = q^k \sum_i (1/m + sv_i)^k$, so

$$l_n'(0) = 0, \qquad l_n''(0)/l_n(0) = mk(k-1).$$

Therefore

$$\langle L'', L\rangle \ge mD^2/8, \qquad \left\|N^{-1}L''\right\|_2 \le mD/4, \qquad \|L\|_2 \le D/4. \tag{30}$$

Primes here denote s derivatives at zero. The inverse is restricted to positive Hermite degrees. A nonzero g gives $D > 0$ for every $q > 0$.

6.2 Higher-product interpolation and analytic completion

For complex z with $|z| \le r$ set $z_i = 1/m + zv_i$. For a subset H of size j, define

$$P_H = q\sum_{i\in H} |z_i|, \qquad \overline{P}_j = cqj/m.$$

Cauchy–Schwarz on the coordinates of v and (27) give

$$\frac{P_H}{\overline{P}_j} \le \frac{1 + mr/\sqrt{j}}{c} \le \rho = \frac{c+3}{4c} < 1, \tag{31}$$

$$\overline{P}_j^{\,j} < q/m, \qquad \overline{P}_j < 1. \tag{32}$$

If $P_H > 0$, normalize $w_i = |z_i|/\sum_{k\in H}|z_k|$. Absolute coefficient majorization in (28) reduces the subset to a real Gaussian product at total signal P_H and these weights, with inputs g_{abs}. This is the L^2 function with Hermite coefficients $|a_n|$; its norm is at most δ, but it need not be bounded. If $P_H = 0$, the positive-degree subset is zero and no normalized weights are needed.

For j independent smoothed Gaussian inputs and normalized nonnegative weights w, conditioning their product on the weighted Gaussian sum is an L^2 contraction. Hence, for $h \in L^2(\varphi)$,

$$\|B_j(P, w; h, \dots, h)\|_2 \le \left\|T_{\sqrt{P}}h\right\|_2^j. \tag{33}$$

Prove this first for finite Hermite sums by the coefficient formula. Independence gives the product L^2 norm. The same multilinear bound applied to a telescoping difference of products extends it to arbitrary L^2 inputs; no product on a single one-dimensional L^2 space is used.

All j factors can be retained in a spectral Hölder interpolation. For nonnegative x_n,

$$\sum_n x_n P^n \le (\sum_n x_n)^{(j-1)/j} (\sum_n x_n P^{jn})^{1/j}.$$

With x_n the squared Hermite coefficients of h and $\|h\|_2 \le \delta$, raising to power $j/2$ gives

$$\left\|T_{\sqrt{P}}h\right\|_2^j \le \delta^{j-1} \left\|T_{P^{j/2}}h\right\|_2. \tag{34}$$

By (32), the real product at enlarged signal $\overline{P}_j$ is therefore bounded by

$$\delta^{j-1}\left\|T_{\overline{P}_j^{j/2}}g_{\mathrm{abs}}\right\|_2 \le \delta^{j-1}\left\|T_{\sqrt{q/m}}g\right\|_2 = \delta^{j-1}\|L\|_2/m.$$

The original signal P_H has the additional output attenuation $(P_H/\overline{P}_j)^{n/2}$. Its number-operator cost is at most $\sup_n n\rho^{n/2} \le K$. Sum the finitely many subsets and use $\delta \le 1$ to obtain, throughout the complex disk,

$$\|M_z\|_2 \le S\delta \|L\|_2 , \qquad \|NM_z\|_2 \le KS\delta \|L\|_2 .$$

Strictness in (31) permits a slightly larger disk. Multilinear completion controls finite-input truncation differences, and exponential output attenuation gives uniform coefficient tails, including every fixed number-operator power. The coefficient series thus converge normally in Hilbert space. Their weight exponents are integers, so zeros of complex coordinates cause no branch issue. Both M_z and L_z are Hilbert analytic. Cauchy's estimate and (30) imply

$$\|M\|_2 \le S\delta D/4, \qquad \|NM\|_2 \le KS\delta D/4, \qquad \|M''\|_2 \le S\delta D/(2r^2). \tag{35}$$

The same normal convergence on a sufficiently small complex ball of weights in $\sum z_i = 1$ gives joint parameter regularity. It is not necessary to interpret complex weights as probability laws.

6.3 Entropy curvature and relative absorption

On the positive real part of this weight neighborhood, (29) bounds the density ratio away from zero. Hilbert C^2 regularity, scalar Taylor's formula and the bounded first two derivatives of $(1+u)\log(1+u)$ on this range justify

$$(\mathcal{F}^{(m)}_{q,g})'' = -\langle R'', \log(1+R)\rangle - \int \frac{(R')^2}{1+R}\varphi.$$

Products of L^2 difference quotients converge in L^1; bounded reciprocal multipliers converge in measure and preserve these limits. This also gives continuity of mixed entropy derivatives. Positive Gaussian noise smooths the bounded input factors spatially. Its conditional Gaussian kernel has positive covariance for $q < 1$, so differentiation supplies bounded spatial derivatives through Gaussian polynomial moments. The number-operator bounds above then justify self-adjoint pairings and the Sobolev chain rule.

For $\psi(u) = \log(1+u) - u$, the same fully proved interpolation estimate (16) gives

$$\begin{aligned}\|N\psi(R)\|_2 &\le \left[\frac{\eta}{1-\eta} + \frac{3\eta}{(1-\eta)^2}\right] \|NR\|_2 \\ &\le 7a\delta(1 + KS\delta/4)D.\end{aligned}$$

Discard the nonpositive first-derivative-square term and split the logarithm as $L + M + \psi(R)$. Equations (30) and (35) give the errors

$$\begin{aligned}|\langle L'', M\rangle| &\le (mKS/16)\delta D^2, \\ |\langle L'', \psi(R)\rangle| &\le (7ma/2)\delta D^2, \\ |\langle M'', \log(1+R)\rangle| &\le S\delta D^2/(3r^2).\end{aligned}$$

For the second use $KS\delta/4 \le 1$; for the third use $S\delta \le 1$, $\eta \le 1/4$ and $\|\log(1+R)\|_2 \le (4/3)\|R\|_2 \le 2D/3$. Their sum is at most $A\delta D^2 \le mD^2/16$, against the restoring energy $mD^2/8$. This proves (2) for unit v, and scaling proves it for every tangent vector. Strictness follows from $D > 0$. Joint Hessian continuity and compactness of the unit tangent sphere give a neighborhood for each fixed input and signal, with no common weight-radius conclusion. This completes the theorem.

7 Higher-moment obstructions near a face

The fourth-cumulant sign mechanism has antecedents in Madiman–Nayar–Tkocz [8]; the three-copy noisy face expansion in [5, Section 7] uses that mechanism and the same density functional appearing in the endpoint identity of [2]. Its leading term vanishes under exact fourth-moment matching. We use a higher first differing moment and a differentiated finite jet to prove Theorem 1.1. The argument below supplies all expansions and sign constructions internally.

7.1 A differentiated finite jet at the face

Let $n = 2\ell \geq 6$ and let X have a bounded smooth even Gaussian-relative perturbation, variance one, and Gaussian moments through $n-2$. Fix $q \in (0,1)$ and let $U = X_q$, with density p and $\kappa_n(U) = q^\ell \kappa_n(X) \neq 0$. Set

$$J(a,b) = h(\sqrt{a}\,U_1 + \sqrt{b}\,U_2 + \sqrt{1-a-b}\,U_3).$$

Let J_G replace U_1, U_2 by independent standard Gaussians. It is a function only of $a+b$:

$$J_G(a,b) = h(\sqrt{1-a-b}\,U_3 + \sqrt{a+b}\,G).$$

Define $B_n(p) = \int p^{(n)} \log p$. We claim

$$J(a,b) - J_G(a,b) = -\frac{\kappa_n(U)}{n!}(a^\ell + b^\ell)B_n(p) + R(a,b), \tag{36}$$

$$|\partial^\alpha R(a,b)| \leq C(a+b)^{\ell+1-|\alpha|}, \qquad |\alpha| \leq 2, \tag{37}$$

near the origin in the positive quadrant. Constants depend on the fixed input and signal.

Here is the regularity and finite-jet justification. Put $a = h_1^2$, $b = h_2^2$. In original-input coordinates the gains are $\sqrt{q}h_1, \sqrt{q}h_2, \sqrt{q}\sqrt{1-h_1^2-h_2^2}$ and the independent Gaussian variance is the positive constant $1-q$. The Gaussian-replacement output instead has Gaussian variance $1-q+q(a+b)$. These gain and variance functions are smooth near zero and bounded away from singularity. The proof of Lemma 3.1 applies verbatim to the finite vector of three inputs: kernel derivatives have conditional polynomial ratios, a bounded-input event supplies the quadratic logarithm bound, and conditional Young bounds differentiated entropy integrals by finitely many moments of these inputs and the Gaussian. Higher moments remove their spatial tails uniformly on the gain disk. All required moments exist for the present bounded Gaussian-relative inputs. Thus the gain entropies are smooth. Evenness in the two light gains and (19) give squared-gain $C^{\ell+1}$ regularity, up to axes, using gain derivatives only through $2\ell + 2 = n + 2$.

Finite characteristic-function derivatives, rather than a convergent cumulant series, identify the first discrepancy of a light factor:

$$\chi_U(\sqrt{a}\,\xi) = e^{-a\xi^2/2}\left[1 + \frac{\kappa_n(U)}{n!}(i\xi)^n a^\ell + \text{higher finite gain jet}\right].$$

After multiplying both light factors and the heavy characteristic function, the first density difference from Gaussian replacement is

$$\frac{\kappa_n(U)}{n!}(h_1^n + h_2^n)p^{(n)}.$$

The Fourier signs agree for even n. Changing the heavy gain or the comparison baseline in this correction adds at least two gain degrees; products of two density differences start at degree $2n$.

The first entropy variation is its negative pairing with $1+\log p$. Its mass integral is zero, leaving exactly the coefficient in (36). All lower difference jets vanish. The just-proved $C^{\ell+1}$ regularity gives (37) by Taylor's theorem for derivatives of orders zero, one and two.

On $a=\epsilon+\tau$, $b=\epsilon-\tau$, the Gaussian baseline is constant in τ. Thus

$$\frac{d^2}{d\tau^2}J(\epsilon+\tau,\epsilon-\tau)\bigg|_{\tau=0}=-\frac{2\ell(\ell-1)}{n!}\kappa_n(U)B_n(p)\epsilon^{\ell-2}+O(\epsilon^{\ell-1}). \tag{38}$$

A negative product $\kappa_n(U)B_n(p)$ therefore gives strictly positive curvature at a genuine interior weight for all sufficiently small positive ϵ, for each fixed input. No sign on an entire τ segment is required to disprove concavity.

7.2 The density functional and a negative quadratic direction

For $p=\varphi(1+r)$ with an even bounded smooth relative perturbation, $\|r\|_\infty<1/2$, integration by parts and the logarithm expansion give

$$\begin{aligned}B_n(p)&=\mathbb{E}_\varphi[\mathrm{He}_n\, r]+Q_n(r)+\text{cubic remainder},\\ Q_n(h)&=\int(\varphi h)^{(n)}h-\tfrac12\mathbb{E}_\varphi[\mathrm{He}_n\, h^2].\end{aligned} \tag{39}$$

Indeed $\int\varphi^{(n)}\log\varphi$ and $\int(\varphi r)^{(n)}\log\varphi$ vanish identically because $\log\varphi$ is quadratic and $n\ge3$. The linear and quadratic terms left by $\log(1+r)$ are exactly those displayed. For the later family $r=\alpha H+\beta\alpha^2W$ with fixed bounded smooth OU images H,W, the remainder is rigorously $O(\alpha^3)$: their spatial derivatives are bounded, $(\varphi H)^{(n)}/\varphi$ and the analogous W expression have polynomial growth, and differentiated logarithms have Gaussian-polynomial integral majorants for small amplitude.

Take the ideal polynomial direction e_d with $d=n+2$. For even n,

$$\frac{(\varphi e_d)^{(n)}}{\varphi}=\sqrt{(d+n)!/d!}\,e_{d+n}.$$

Its pairing with e_d is zero. The Hermite product formula $\mathrm{He}_d^2=\sum_j\binom{d}{j}^2 j!\,\mathrm{He}_{2d-2j}$ gives

$$\mathbb{E}_\varphi[\mathrm{He}_n\, e_d^2]=\frac{n!d!}{(d-\ell)!(\ell!)^2}=:c_{n,d}>0.$$

Consequently

$$Q_n(T_{\sqrt q}e_d)=-\tfrac12 q^d c_{n,d}<0. \tag{40}$$

This unbounded polynomial selects an approximation direction; it is never used as a bounded relative density.

7.3 Compact moment corrections and continuity

Let $I=\{0,2,\dots,n\}$. Choose a nonnegative even smooth compact cutoff χ positive on a nonempty open interval, and put

$$\psi_j=\chi\,\mathrm{He}_j,\qquad \Gamma_{ij}=\int\chi\,\mathrm{He}_i\,\mathrm{He}_j\,\varphi\quad(i,j\in I).$$

The matrix Γ is positive definite, since a nonzero polynomial cannot vanish on that interval, and is therefore invertible. Choose even smooth cutoffs χ_R equal to one on $[-R, R]$, zero outside $[-2R, 2R]$, and between zero and one. For $u_{R,i} = \mathbb{E}_\varphi[\mathrm{He}_i\, \chi_R e_d]$ define

$$v_R = \chi_R e_d - \sum_{j \in I} (\Gamma^{-1} u_R)_j \psi_j.$$

Then $\mathbb{E}_\varphi \mathrm{He}_i\, v_R = 0$ exactly for every $i \in I$. Also $v_R \to e_d$ in $L^2(\varphi)$: cutoff convergence follows by dominated convergence and $u_R \to 0$ by orthogonality and $d > n$. Every v_R is even, bounded, smooth and compactly supported.

For fixed $q < 1$, the map $v \mapsto Q_n(T_{\sqrt{q}} v)$ is a continuous quadratic form on $L^2(\varphi)$. To verify the needed continuity explicitly, the raising operator $A_n h = (\varphi h)^{(n)}/\varphi$ sends e_k to $\sqrt{(k+n)!/k!}\, e_{k+n}$ for even n. Multiplication by He_n is a finite sum of Hermite shifts with coefficients bounded by $C_n(1+k)^{n/2}$. After either operator is composed with $T_{\sqrt{q}}$, its exponentially attenuated coefficients are bounded, so each composition is bounded from L^2 to L^2. The pairings with $T_{\sqrt{q}} v$ in (39) are thus continuous. This does not require uniform supremum or derivative bounds on the approximating cutoff polynomials.

By (40), freeze a sufficiently large R so that $v = v_R$ has $Q := Q_n(T_{\sqrt{q}} v) < 0$. Let $e^{(n)}$ be the coordinate vector indexed by n in I, and freeze

$$w = \sum_{j \in I} (\Gamma^{-1} e^{(n)})_j \psi_j.$$

It satisfies $\mathbb{E}_\varphi \mathrm{He}_i\, w = 0$ for $i < n$ and $\mathbb{E}_\varphi \mathrm{He}_n\, w = 1$. The two directions may depend on fixed q, n, but not on the amplitude or face distance.

7.4 Actual densities and the order of choices

Choose $\beta > 0$ with $\beta q^\ell < -Q/2$. For positive α tending to zero put

$$g_\alpha = \alpha v + \beta \alpha^2 w, \qquad f_\alpha = \varphi(1 + g_\alpha).$$

All moments through $n-2$ are exactly Gaussian. The first differing cumulant is exactly

$$\kappa_n(X_\alpha) = \mathbb{E}\, \mathrm{He}_n(X_\alpha) = \beta \alpha^2 > 0.$$

Here $\mathbb{E}_\varphi \mathrm{He}_n\, w = 1$ is unnormalized; no factor $\sqrt{n!}$ is missing. The smoothed law has relative function $r_\alpha = \alpha H + \beta \alpha^2 W$, with $H = T_{\sqrt{q}} v$ and $W = T_{\sqrt{q}} w$. Equation (39) therefore gives

$$B_n(p_\alpha) = \alpha^2 (Q + \beta q^\ell) + O(\alpha^3) < 0$$

for small positive amplitude, whereas $\kappa_n(U_\alpha) = q^\ell \beta \alpha^2 > 0$.

Because v, w are fixed compact smooth functions, $\|g_\alpha\|_{C_b^r} \to 0$ for every fixed r. For small amplitude, $1 + g_\alpha \geq 1/2$ and

$$(\log f_\alpha)'' = -1 + \frac{g_\alpha''}{1+g_\alpha} - \frac{(g_\alpha')^2}{(1+g_\alpha)^2} \leq -1 + 2\, \|g_\alpha''\|_\infty \leq -1/2.$$

Thus these are positive smooth strongly log-concave even densities, with mass and variance exactly one. Exact Hermite matching is equivalent to ordinary moment matching because the Hermites form a triangular polynomial basis; odd moments match by symmetry.

For each fixed sufficiently small amplitude, (38) now has strictly positive leading coefficient. Choose a face distance $\epsilon(\alpha) > 0$ small enough to dominate its remainder, and also less than α and $1/4$. Taking $n = 2K + 2$ and a sequence $\alpha_j \downarrow 0$ gives Theorem 1.1, with $\epsilon_j \to 0$ (and a decreasing subsequence if necessary). The order is fixed q, K, then compact directions and β, then amplitude, then face distance. No uniform-in-amplitude remainder or fixed face weight is asserted. The equality in law between mixing copies of $U = X_q$ and adding Gaussian variance $1 - q$ after mixing copies of X identifies this J with the three-copy entropy in the theorem.

Corollary 7.1. *The nonconcavity conclusion of Theorem 1.1 holds for every fixed number of copies $m \geq 3$, at a true interior simplex point, with the same prescribed finite moment matching and Gaussian-relative convergence. The weights may depend on the input.*

Proof. For $m > 3$, embed the three-copy point with the remaining weights zero. Its positive tangent curvature is unchanged. The fixed positive Gaussian noise and the finite-vector gain argument used above give continuous weight Hessians up to these axes by the even-square formula. Therefore its positive curvature persists after sufficiently small positive extra weights are added, scaling the first three weights by their remaining total mass. All m weights are then positive. Apply this continuity argument separately to each input in the constructed sequence; no common extra-weight size is needed. □

8 Restoration regimes and moment topology

Theorem B.1 supplies a concrete distinction between the two neighborhoods. Its laws match any fixed finite set of Gaussian moments exactly and approach the Gaussian in total variation and every fixed absolute moment, while their positive-signal smoothings retain exponential tails and an unbounded Gaussian-relative density ratio. This makes the moment and relative-density regimes mathematically complementary.

9 Sharp barycentric thresholds for every number of copies

9.1 Results

Let γ be the standard Gaussian law on $\mathbb{R}$. Write $e_n = \mathrm{He}_n/\sqrt{n!}$ for the normalized probabilists' Hermites, $T_\rho e_n = \rho^n e_n$ for the Ornstein–Uhlenbeck operator, and $\mathcal{N} e_n = n e_n$. All unmarked L^p norms below are for γ. Call g admissible if g is a real even bounded function with $\int g\, d\gamma = \int x^2 g\, d\gamma = 0$. For $\|g\|_\infty < 1$, let X_i be independent with law $(1+g)\gamma$, and let $G \sim \gamma$ be independent. For positive weights $w = (w_1, \dots, w_m)$ with $\sum_i w_i = 1$, define

$$F_{q,g}(w) = h\left(\sqrt{q}\sum_{i=1}^{m}\sqrt{w_i}\, X_i + \sqrt{1-q}\, G\right), \qquad 0 < q < 1.$$

Here h is differential entropy. Write $u_m = (1/m, \dots, 1/m)$.

Define

$$\lambda_a = \frac{a}{1+\sqrt{1-a^2}}, \qquad \tau_a(b) = \frac{\sqrt{1-a^2} - 1 + b}{\sqrt{1-a^2} + 1 - b} \quad (0 < a \leq b < 1).$$

There is a unique $q_m \in (0, 1)$ such that

$$\tau_{q_m/m}(q_m)^m = \lambda_{q_m/m}. \tag{41}$$

In particular $q_3 = 0.65767078\ldots$ and $q_4 = 0.71210929\ldots$.

Theorem 9.1. *Fix $m \geq 3$. For every $0 < Q < q_m$ there are $\delta, \rho > 0$ such that for every admissible g of norm at most δ, every $q \in (0, Q]$, and every weight w with $\|w - u_m\|_2 \leq \rho$, the Hessian of $F_{q,g}$ on the tangent space is negative definite unless $g = 0$ almost everywhere.*

Theorem 9.2. *One pair $\delta, \rho > 0$ works for all $0 < q \leq q_m$ in the following radial assertion. For every tangent unit vector v, $s \mapsto F_{q,g}(u_m + sv)$ has strictly negative second derivative on $0 \leq s \leq \rho$ whenever $0 < \|g\|_\infty \leq \delta$. In particular the Hessian at u_m is negative definite also at $q = q_m$. For any fixed $q \in (q_m, 1)$ and any prescribed finite Gaussian moment matching order there are smooth strongly log-concave laws $(1 + g_n)\gamma$ matching those moments, with $\|g_n\|_{C_b^k} \to 0$ for every fixed k, and with positive definite tangent Hessian at u_m.*

Theorem 9.3. *For every fixed $m \geq 3$ and $0 < \eta < 1/m$, there is a positive relative-density radius on which the Hessian is strictly negative in every nonzero tangent direction throughout $\min_i w_i \geq \eta$, for every non-Gaussian admissible input and all $0 < q \leq \eta^{3/2}/(256\sqrt{m})$.*

Theorem 9.3 answers the fixed-interior existence question in [5, Section 9.3]. Luo [7] supplies the oscillator mechanism and the 4/7 obstruction in the two-copy problem. The Evidence Press candidate [5] develops sharp bounded-function oscillator interpolation and a three-copy obstruction near a face. The theorem of Section 6 gives the multicopy Hessian at exactly equal weights for $Q < m^{-1/(m-1)}$, and moment-matched failures near a face. The L^p oscillator estimate below controls the loss of directional coercivity relative to the largest weight off the barycentre. The high-signal proof uses a different obstruction at exactly equal weights, from a coefficient of a positive Gaussian quadratic exponential.

For the threshold algebra, set $a = q/m$ and $\lambda = \lambda_a$. Then

$$\tau_a(ja) = \lambda \frac{j - \lambda}{1 - j\lambda} \qquad (1 \leq j \leq m).$$

For $j = m$, $\tau_a(q)^m/\lambda$ is strictly increasing from zero as q increases, and tends to a value greater than one as $q \uparrow 1$, since $\tau_a(1) = 1$. This gives the unique threshold. For fixed a, $j \mapsto j \log \tau_a(ja)$ is strictly convex for $1 \leq j \leq m$. Indeed, writing $d = \sqrt{1 - a^2}$, differentiation gives

$$\partial_b \log \tau_a(b) = \frac{2d}{d^2 - (1 - b)^2}, \qquad \partial_b(b\partial_b \log \tau_a(b)) > 0 \quad (b > a).$$

The derivative of $j \log \tau_a(ja)$ is the sum of two strictly increasing functions of j. Thus $\tau_a(ja)^j < \lambda_a$ for all $j \geq 2$ when $q < q_m$. At equality it is strict for $2 \leq j < m$. When $q > q_m$, for $j < m$ it is strictly smaller than $\tau_a(q)^m$, including $j = 1$.

9.2 Entropy and spectral comparison

Write $R_w = L_w + M_w$ for the output density ratio minus one relative to γ. Conditioning Gaussian inputs on their normalized output gives a sum over nonempty subsets $H \subseteq [m]$, with the singletons forming $L_w = \sum_i T_{\sqrt{qw_i}} g$. If $g = \sum_n a_n e_n$, the degree $n = \sum_{i \in H} n_i$ contribution of the subset H to the ratio is

$$\left(\frac{n!}{\prod_{i \in H} n_i!} \right)^{1/2} q^{n/2} \prod_{i \in H} w_i^{n_i/2} a_{n_i}. \tag{42}$$

Every active n_i is even and at least four. This formula follows from the Gaussian generating functions and holds by finite approximation for L^2 inputs. At total signal $b = q \sum_{i \in H} w_i$ the subset is the conditional expectation of a product of independent $T_{\sqrt{b}} g$ factors, onto the sum with weights

$w_i/\sum_{k\in H} w_k$. In particular its norm is at most $\|T_{\sqrt{b}}g\|_2^{|H|}$. The conditional projection sends tensor total number to $\mathcal{N}$, as is seen in the same coefficient formula. For complex weights, replace each weight by its modulus and rotate its input Hermite coefficients by the corresponding unit phases $e^{in_i \arg w_i/2}$. This preserves the smoothed input norms and proves the same bound at $b = q\sum|w_i|$. An increase of b adds exponential coefficient damping at the original signal. It pays for arbitrary fixed number powers and gives normal Hilbert convergence of the analytic series on any smaller complex weight disk with strict margin.

For a unit tangent direction v, primes below mean derivatives in this direction at a real w. For degree $n = 2k \geq 4$ the multiplier of a_n in L is $q^k\sum_i w_i^k$, and its second derivative is $q^k k(k-1)\sum_i v_i^2 w_i^{k-2}$. Consequently $E = \langle L'', L\rangle > 0$ if $g \neq 0$. On fixed positive weight regions one has

$$\|\mathcal{N}^{-1}L''\|_2 \leq CE^{1/2}, \qquad E \geq c\|\mathcal{N}T_{\sqrt{a}}g\|_2^2 \tag{43}$$

whenever $a \leq q\sqrt{(\min_i w_i)(\max_i w_i)}$. The first estimate follows from the ratio of the second multiplier to the original multiplier being at most Cn^2. The second follows by keeping a largest weight in the original multiplier and bounding the second multiplier from below by its value using the smallest weight. On a ray $w = u_m + sv$, $s \geq 0$, a largest weight has $v_i = \max_j v_j \geq [m(m-1)]^{-1/2}$. This gives the stronger $E \geq c_m\|\mathcal{N}L\|_2^2$ there.

Put $\delta = \|g\|_\infty$. For real weights R and its first two spatial derivatives have uniform sup norm $O(\delta)$ for $q \leq Q < 1$. Indeed the conditional Gaussian input vector has covariance $I - aa^\top$, $a_i = \sqrt{qw_i}$, bounded below by $1-Q$. Differentiate its density in the conditional mean ax and use the bound on the input relative product minus one. This also gives smoothness in the spatial variable. In strictly positive weight regions the parameter derivatives up to order two exist in L^2; one may differentiate the same conditional Gaussian density, whose derivatives are bounded in absolute expectation by a polynomial in the output variable uniformly on compact parameter sets. It follows, by scalar differentiation of entropy and boundedness of the real density ratio away from zero for small δ, that

$$F'' = -\langle R'', R + \psi(R)\rangle - \int \frac{(R')^2}{1+R}\,d\gamma, \qquad \psi(z) = \log(1+z) - z. \tag{44}$$

For sufficiently small $\|R\|_\infty$,

$$\|\mathcal{N}\psi(R)\|_2 \leq C\|R\|_\infty\|\mathcal{N}R\|_2. \tag{45}$$

Indeed $\mathcal{N} = -\partial_x^2 + x\partial_x$, and Gaussian integration by parts gives

$$\|R_x^2\|_2^2 = \int RR_x^2(xR_x - 3R_{xx})\,d\gamma \leq 3\|R\|_\infty\|R_x^2\|_2\|\mathcal{N}R\|_2,$$

using $\|R_{xx}\|_2 \leq \|\mathcal{N}R\|_2$ from Hermite expansion. The chain rule gives (45).

9.3 Oscillator interpolation

We give the strict and weighted oscillator estimates used below. The Lebesgue oscillator basis and its Gaussian Gram calculation are due to the approach in [7, 5]. For $0 < a < 1$, set $d = \sqrt{1-a^2}$, $B = d^{-1}$ and

$$H_k(x) = c_B e^{-Bx^2/2}\frac{\mathrm{He}_k(\sqrt{2B}\,x)}{\sqrt{k!}},$$

where the common constant c_B gives an orthonormal basis of $L^2(dx)$. Let $S_k(b) = \|T_{\sqrt{b}}H_k\|_2^2$. For two standard Gaussians correlated by b, direct integration of the two Hermite generating functions gives, up to a positive factor depending on a, b,

$$\exp\left\{\frac{\nu(z+w)^2 - \tau(z-w)^2}{4}\right\}, \quad \nu = \frac{1+b-d}{1+b+d}, \quad \tau = \frac{d-1+b}{d+1-b}.$$

At $b = a$ one has $\nu = \tau = \lambda_a$. Thus the images at a are orthogonal and $S_k(a) = C_a \lambda_a^k$. To extract the diagonal for other b, substitute $z = \sqrt{t}\, W$, $w = \sqrt{t}\, \overline{W}$, where W is a standard circular Gaussian with $\mathbb{E} W^j \overline{W}^k = k!\, \mathbf{1}_{j=k}$, and integrate. The resulting diagonal generating function is a constant times $[(1 - t\nu)(1 - t\tau)]^{-1/2}$. In particular, for $b > a$, since $0 < \nu < \tau < 1$,

$$S_k(b) \le C(k+1)^{-1/2} \tau_a(b)^k. \tag{46}$$

Indeed the binomial coefficients for $(1-z)^{-1/2}$ are $O((k+1)^{-1/2})$; convolution with the geometrically damped coefficients of the other factor preserves that bound. These constants are uniform on compact regions where $a, b - a$ are bounded below and $b < 1$ is bounded away from one. For $b \le a$ one has $|\tau| \le \nu$ and the same generating function gives $S_k(b) \le C\nu^k$.

For bounded g, the oscillator coefficients $c_k = \int g H_k \, dx$ satisfy

$$|c_k| \le C \min\left\{ \delta (k+1)^{1/4}, D \lambda_a^{-k/2} \right\}, \qquad D = \|T_{\sqrt{a}} g\|_2. \tag{47}$$

For the first bound use weighted Cauchy–Schwarz with weight $1 + x^2/(k+1)$, and $\int x^2 |H_k|^2 dx = O(k+1)$ by the oscillator recurrence. For the second use orthogonality of the smoothed basis at a. One may first apply this to g cut off on a large interval. It then follows for bounded g by $L^2(\gamma)$ convergence and integrability of H_k. The series at $b < 1$ represents the smoothed limit, since the coefficients have polynomial growth and the smoothed modes have geometric decay. In particular, (46) and interpolation of the two bounds in (47) give

$$\|T_{\sqrt{b}} g\|_2 \le C \delta^{1-p} D^p, \qquad 0 < p < \frac{\log \tau_a(b)}{\log \lambda_a}. \tag{48}$$

We also need the $L^P(\gamma)$ version for $P > 2$. For this, integrate the Mehler sum of the two basis functions against independent $N(\sqrt{b}\, x, 1 - b)$ variables. Writing

$$D_\pm(z) = 1 + B(1-b) \ \pm \big(1 - B(1-b)\big) z,$$

this gives

$$\sum_{k \ge 0} z^k |T_{\sqrt{b}} H_k(x)|^2 = \frac{C_B}{\sqrt{D_+(z) D_-(z)}} \exp\left(\frac{bB(z-1)x^2}{D_+(z)} \right). \tag{49}$$

Indeed the sum and difference of the two integration variables diagonalize the quadratic in Mehler's formula. The assertion first holds for $0 < z < 1$, hence as a Taylor identity, and the right side extends the series to $0 < z < 1/\tau_a(b)$ when $b \ge a$. Taking $z = 1/w_0 > 1$, each mode has L^P norm at most $C w_0^{k/2}$ if $w_0 > \tau_a(b)$ and the exponential to power $P/2$ is integrable against γ. The latter condition is $w_0 > \sigma_a(b, P)$, where

$$\sigma_a(b, P) = \frac{B[1 + (P-1)b] - 1}{B[1 + (P-1)b] + 1}.$$

Combining the bound for each mode with (47) and summing proves

$$\|T_{\sqrt{b}} g\|_P \le C \delta^{1-p} D^p, \qquad 0 < p < \frac{\log \max\{\tau_a(b), \sigma_a(b, P)\}}{\log \lambda_a}. \tag{50}$$

All assertions with a strict margin are uniform on compact parameter sets.

The important feature is that at $b = a$ and $P = 2$, both $\tau_a(a)$ and $\sigma_a(a, 2)$ equal λ_a. For $P > 2$ the latter is larger, and

$$\partial_P \log \sigma_a(a, P)|_{P=2} = \sqrt{\frac{1-a}{1+a}} =: z_a, \qquad \log(1/\lambda_a) = 2\,\mathrm{artanh}(z_a) > 2z_a.$$

Therefore for $P > 2$ sufficiently close to 2, one may take p in (50) with $pP/2 > 1$. This strict inequality persists with $b > a$ sufficiently close to a. The choices are uniform for a in a compact subinterval of $(0, 1)$, by the strict derivative inequality.

Finally the radial endpoint needs a number-weighted bound without enlarging b. Write $X = \|(\mathcal{N}+1)T_{\sqrt{a}}g\|_2$. The second estimate in (47) improves to

$$|c_k| \le CX(k+1)^{-1}\lambda_a^{-k/2}.$$

Indeed

$$\|(\mathcal{N}+1)^{-1}T_{\sqrt{a}}H_k\|_2^2 = \int_0^1 (-\log v)S_k(av)\,dv \le C(k+1)^{-2}\lambda_a^k.$$

For the last estimate the diagonal formula above gives $S_k(av) \le C\lambda_a^k e^{-ck(1-v)}$, by monotonicity and a positive derivative of ν on $0 \le av \le a$. Similarly (46), with b tilted to $b\exp(c/(k+1)) < 1$ for a small fixed c, gives

$$\|\mathcal{N}^j T_{\sqrt{b}}H_k\|_2 \le C_j(k+1)^{j-1/4}\tau_a(b)^{k/2}$$

for each fixed integer $j \ge 0$. This follows from $n^{2j} \le C_j(k+1)^{2j}e^{cn/(k+1)}$ at Hermite degree n, and the bounded logarithmic derivative of τ_a on the parameter range. Split the oscillator series where $(K+1)^{5/4}\lambda_a^{K/2}$ is comparable to X/δ. This is possible with constants on the compact parameter range, unless X/δ is bounded below, in which case use $K = 0$ and the uniform sup-norm coefficient bound. Summing the two geometric tails gives $\|\mathcal{N}^j T_{\sqrt{b}}g\|_2 \le C\delta(K+1)^j\tau_a(b)^{K/2}$ for $j = 0, 1$. In particular, for an integer $l \ge 2$,

$$\|\mathcal{N}T_{\sqrt{b}}g\|_2\,\|T_{\sqrt{b}}g\|_2^{l-1} \le C\delta^{l-1}X \qquad \text{if } \tau_a(b)^l \le \lambda_a. \tag{51}$$

Here $\delta(K+1)\lambda_a^{K/2} \le CX$ by the choice of the split. The upper and lower geometric ratios used in the split are uniform when $b - a > 0$ is bounded below and both levels stay in a compact subset of $(0, 1)$. If $g = 0$, all these estimates are immediate.

9.4 Strict concavity on real weight regions

9.4.1 Every fixed region away from the faces

Fix w with $\min_i w_i \ge \eta$ and put $w_* = \max_i w_i$, $a = q\sqrt{\eta w_*}$, $D = \|T_{\sqrt{a}}g\|_2$, and $p_0 = \sqrt{\eta}/2$. The inverse $T_{1/\sqrt{p_0}}$ in the following bounds means coefficientwise multiplication by $p_0^{-n/2}$ on functions for which the resulting series is in L^2. On a complex weight disk of fixed small radius along a unit tangent direction, the total signal of any subset is at most $2q$. After that inverse multiplier, use the enlarged signal $b = 4q/p_0$. For $q \le \eta^{3/2}/(256\sqrt{m})$ one has $b^2 \le a/4$ and $b < 1$, since $w_* \ge 1/m$. For $j \ge 2$, spectral Hölder gives

$$\|T_{\sqrt{b}}g\|_2^j \le \delta^{j-1}\|T_{b^{j/2}}g\|_2. \tag{52}$$

This follows by Hölder applied to the sum of squared Hermite coefficients times b^n. Thus every mixed subset is $O(\delta D)$ after the inverse multiplier. The remaining attenuation from the enlargement justifies normal analytic convergence and Cauchy's estimate. In particular

$$\|T_{1/\sqrt{p_0}}M\|_2 + \|T_{1/\sqrt{p_0}}M''\|_2 \le C\delta D. \tag{53}$$

Constants are uniform on the trimmed weight region.

Since $p_0 w_* \le \sqrt{\eta w_*}$, the inner product $\langle M'', L\rangle$ is $O(\delta D^2)$ by transferring the attenuation to L. Also $\|\mathcal{N}M\|_2 \le C\delta D$ by the attenuation in (53), so $|\langle L'', M\rangle| \le C\delta D E^{1/2}$. The M'' and M pairing is $O(\delta^2 D^2)$.

For the nonlinear logarithm one can use L^3. Gaussian hypercontractivity [11] $T_{1/\sqrt{2}} : L^2 \to L^3$ and (53) give $\|M\|_3 + \|M''\|_3 \le C\delta D$. For L, L'', enlarge the individual signal qw_i to $3qw_*$, using part of that enlargement to pay the number powers from the derivatives and part for hypercontractivity. Since $(3qw_*)^{3/2} \le a$, the same spectral interpolation as in (52) gives

$$\|L\|_3 + \|L''\|_3 \le C\delta^{1/3} D^{2/3}.$$

As $D \le \delta \le 1$, the same bound holds for R, R''. Using $|\psi(R)| \le C|R|^2$ gives $|\langle R'', \psi(R)\rangle| \le C\delta D^2$. In (44), the main pairing is E and $D \le CE^{1/2}$ by (43). The indicated errors are at most $C\delta E$. This proves Theorem 9.3.

9.4.2 A common weight ball below the sharp boundary

Fix $Q < q_m$. For all sufficiently small q, the preceding trimmed argument applies on a fixed ball about u_m. We therefore consider q in a compact interval $[q_0, Q]$ with $q_0 > 0$. Take the common weight ball of radius ξ and fix $a = q(1/m - \xi)$ there. The ball and its comparison level are reduced together when ξ is decreased below. Put $D = \|T_{\sqrt{a}} g\|_2$. The strict subset inequalities at the barycentre allow the following simultaneous choices. Take $p_0 < 1$ close to one, shrink the ball, and take a uniform complex disk radius about each real weight in the ball. After the inverse multiplier $T_{1/\sqrt{p_0}}$, every size-j subset on a disk has signal below an enlarged $b_j > a$ with $\tau_a(b_j)^j < \lambda_a$. Also $p_0 \max_i w_i \le a/q$. These choices follow by strictness and continuity on $[q_0, Q]$ and the finite number of subsets. Any positive separation between the original and enlarged levels is chosen uniformly relative to q. Now use (48) with $p = 1/j$. The subset norm and Cauchy's estimate give (53) again. The same three mixed pairings as above are bounded by $C\delta E$.

For the logarithm choose $P > 2$ sufficiently close to 2 that both $p_0(P-1) < 1$ and the strict oscillator estimate (50) has an exponent p_1 with $p_1 P/2 > 1$ at $b = a$, uniformly on the compact parameter set. The construction of P, p_1 follows from the uniform strict derivative inequality after (50). Shrink the weight ball further so that one can choose $b > \max_i qw_i$ close to a with the same estimate. The ratio qw_i/b is uniformly below one. After factoring through $T_{\sqrt{b}}$, spatial derivatives through order two are bounded as operators $L^P(\gamma) \to L^P(\gamma)$. Indeed differentiating a fixed residual Ornstein–Uhlenbeck kernel introduces only fixed Gaussian Hermite factors in the conditional integral, and conditional Hölder followed by integration in the output bounds the norm. Hence the L^P norms of L, L_x, L_{xx} are at most $C\delta^{1-p_1} D^{p_1}$. For M and its first two spatial derivatives, (53) gives the stronger $C\delta D$ bound by splitting the reserved multiplier $T_{\sqrt{p_0}}$ between hypercontractivity and a fixed residual kernel. For R and these derivatives the L^∞ bounds are $C\delta$.

Choose $U > 4$ close enough to four that $2Pp_1/U > 1$, with $P < U$. Interpolating between L^P and L^∞ gives

$$\|\partial_x^j R\|_U \le C\delta^{1-p_1 P/U} D^{p_1 P/U}, \qquad j = 0, 1, 2.$$

Each term in the number chain rule $\mathcal{N}\psi(R) = \psi'(R)(-R_{xx} + xR_x) - \psi''(R)R_x^2$ is controlled in L^2 by a product of two such bounds. For the factor x use its finite Gaussian L^A norm with $1/A + 2/U = 1/2$; when there is no x, use that γ is a probability law. Since $|\psi'(R)| \le C|R|$, $|\psi''(R)| \le C$, and $D \le \delta$, this proves

$$\|\psi(R)\|_2 + \|\mathcal{N}\psi(R)\|_2 \le C\delta D.$$

Pair $\mathcal{N}^{-1}L''$ with $\mathcal{N}\psi(R)$, and use the direct L^2 estimate on M'' for its pairing with $\psi(R)$. All errors in (44) are $O(\delta E)$. Choosing a common small δ proves Theorem 9.1.

9.4.3 Radial concavity at the boundary

For q bounded away from zero and at most q_m, take $a = q \max_i w_i$ and a sufficiently small common weight ball. For a subset of size $j \geq 2$ at real weights, use $b = q \sum_{i \in H} w_i$. For proper subsets the strict margins give $\tau_a(b)^j < \lambda_a$ by compactness. For the full subset $b = q$ and $a \geq q/m$. The function $\tau_a(q)$ decreases with a, while λ_a increases, so $\tau_a(q)^m \leq \lambda_a$, including at the threshold. In this common ball $b - a$ is bounded below for these subsets, so (51) applies uniformly.

Let $D_* = \|\mathcal{N}L\|_2$. We have $\|(\mathcal{N}+1)T_{\sqrt{a}}g\| \leq CD_*$ by the nonnegative linear multipliers and absence of low degrees. By conditional product contraction, commutation of $\mathcal{N}$ with total number and (51),

$$\|\mathcal{N}M\|_2 + \|\mathcal{N}^{-1}M''\|_2 \leq C\delta D_*.$$

For the derivative estimate, differentiate the coefficient (42) in the real weights. The resulting factor is a polynomial in the active degrees bounded by $C(\sum n_i)^2$. After division by the output degree, its cost in the tensor norm is at most the total number. The total number of a tensor product is the sum of the numbers in the single factors, which is exactly the product of norms appearing in (51), up to a fixed combinatorial factor. All manipulations can first be performed on finite series. Because b stays below one, the smoothed number norms justify their limits and the same real derivatives.

Along a ray the linear energy has $E \geq cD_*^2$. The estimate (45) and the bound on $\mathcal{N}M$ give $\|\mathcal{N}\psi(R)\|_2 \leq C\delta D_*$. Pair the two second derivatives through $\mathcal{N}^{-1}$ in (44). All errors are $O(\delta D_*^2)$; the first derivative squared term has a nonpositive sign. Small q is covered by Theorem 9.3. Choosing the common radius and relative-density bound proves the positive part of Theorem 9.2.

9.5 Positive barycentric curvature above the threshold

Fix $q_m < q < 1$. In this section put $a = q/m$, $B = (1-a^2)^{-1/2}$, and abbreviate $\lambda = \lambda_a$, $\tau = \tau_a(q)$, and $\nu = (1+q-\sqrt{1-a^2})/(1+q+\sqrt{1-a^2})$. We use the even oscillator modes H_n of Section 3 as $n \to \infty$ through even integers. We first compute the leading m-factor term of their output density.

Let $B_{m,n}(w)$ denote the conditional expectation of $\prod_{i=1}^m H_n(X_i)$ for independent standard Gaussian inputs and the given noisy output. Its squared $L^2(\gamma)$ norm is the expectation of the product of all $2m$ copies of H_n on two input blocks drawn independently conditionally on one output. Writing $d_w = (\sqrt{w_1}, \dots, \sqrt{w_m})^\top$, their joint covariance has diagonal blocks I and cross blocks $q d_w d_w^\top$. Denote it by Σ_w. Integrating the oscillator generating functions gives

$$\|B_{m,n}(w)\|_2^2 = C_{m,q}(n!)^m [z_1^n \cdots z_{2m}^n] \exp\left(\tfrac{1}{2} z^\top Q_w z\right), \qquad Q_w = 2B\Sigma_w(I + B\Sigma_w)^{-1} - I. \tag{54}$$

The prefactor is positive and independent of w, since the eigenvalues of Σ_w are $1+q, 1-q, 1, \dots, 1$. The generating formula follows by integrating each $e^{-Bx^2/2}\exp(\sqrt{2B}xz - z^2/2)$; the $\sqrt{n!}$ in each extracted normalized coefficient accounts for $(n!)^m$.

Replace the variables of the first block by iz_j and the second block by $-iz_j$. The target coefficient is unchanged by these phases. The resulting real symmetric matrix $\widetilde{Q}_w$ has eigenvalue τ in direction (d_w, d_w), eigenvalue $-\nu$ in direction $(d_w, -d_w)$, and eigenvalue $-r_0$ on their orthogonal complement, where $r_0 = (B-1)/(B+1) = \lambda^2$. Thus the diagonal blocks are

$$-r_0 I + (r_0 + (\tau - \nu)/2) d_w d_w^\top$$

and the cross blocks are $(\tau+\nu)d_w d_w^\top/2$. At the barycentre all entries are strictly positive. Only the diagonal entries require checking. The identities

$$\tau = \lambda\frac{m-\lambda}{1-m\lambda}, \qquad \nu = \lambda\frac{m+\lambda}{1+m\lambda}$$

give

$$(\tau-\nu)/2 = \frac{\lambda^2(m^2-1)}{1-m^2\lambda^2} > (m-1)r_0.$$

Positivity therefore holds on a neighborhood of the barycentre. Every row sum of $\widetilde{Q}_{u_m}$ is τ.

For clarity we analyze the coefficient and its derivatives. For $x \in \mathbb{R}^{2m}$, define

$$\Phi_w(x) = \frac{1}{2\tau}\sum_{i,j}\widetilde{Q}_{w,ij}e^{x_i+x_j} - \sum_i x_i.$$

Its gradient vanishes at $w = u_m, x = 0$. Its Hessian in x there is positive definite, since the diagonal terms contribute positive squares and all other terms contribute nonnegative squares. The same holds nearby. Let $x(w)$ be the smooth critical point near zero given by the implicit function theorem, and write $J(w) = \Phi_w(x(w))$.

Apply Cauchy's coefficient formula in (54) on the radii $\sqrt{n/\tau}\,e^{x_i(w)}$. There are positive smooth numbers $\mu_{ij}(w)$ for $i \le j$, given by $\widetilde{Q}_{w,ij}e^{x_i+x_j}/\tau$ for $i < j$ and half that formula for $i = j$. The critical point equations give $\sum_{j\neq i}\mu_{\min(i,j),\max(i,j)} + 2\mu_{ii} = 1$. The coefficient is

$$(\tau/n)^{mn}e^{nJ(w)}I_n(w)$$

with

$$I_n(w) = \frac{1}{(2\pi)^{2m}}\int_{[-\pi,\pi]^{2m}}\exp\left(n\sum_{i\le j}\mu_{ij}(w)(e^{i(\theta_i+\theta_j)}-1) - in\sum_i\theta_i\right)d\theta.$$

The only maximum-modulus points of the integrand on the torus are the origin and the point with every coordinate π. Indeed the loops require each coordinate to be 0 or π modulo 2π, and the positive ordinary edges require the same choice for all coordinates. Their contributions have the same phase.

We record the derivative version of the Gaussian estimate for this integral. In local coordinates centered at either point the constant phase is one, and the linear part in θ is zero by the degree equations. The local real part is bounded above by $-cn|\theta|^2$ uniformly for w in a fixed small neighborhood. At $\theta = y/\sqrt{n}$ the exponent converges to $-\frac{1}{2}\sum_{i\le j}\mu_{ij}(w)(y_i+y_j)^2$. The same expansion holds for two derivatives in w along the simplex. The derivative of the linear term is zero since the degree equations hold identically in w, so the differentiated integrands on this scale are bounded by a fixed polynomial in $|y|$ times $e^{-c|y|^2}$. Away from the two points, even the differentiated integrals are exponentially small. It follows by dominated convergence that $n^m I_n(w)$ converges in C^2 on a smaller neighborhood to a strictly positive smooth Gaussian integral, with the factor of two from the two points. Uniformity follows from the compact parameter neighborhood and the uniform Taylor remainder. In particular $(\log I_n)'' = O(1)$.

By symmetry $J'(u_m) = 0$ along the simplex. At u_m, differentiating the critical point equations gives

$$J'' = (\Phi_w)''\big|_{x=0} - (\partial_v\nabla_x\Phi_w)^\top(\nabla_x^2\Phi_w)^{-1}(\partial_v\nabla_x\Phi_w)\big|_{w=u_m,x=0} \le (\Phi_w)''\big|_{w=u_m,x=0}.$$

The latter derivative is strictly negative for any unit tangent v, because

$$\frac{1}{2}\mathbf{1}^\top \widetilde{Q}_w \mathbf{1} = -mr_0 + (\tau + r_0)\left(\sum_i \sqrt{w_i}\right)^2$$

and the squared sum has strictly negative second derivative at equal weights along each such direction. Since $J(u_m) = m$, Stirling's formula in (54) gives

$$\|B_{m,n}(u_m)\|_2^2 \asymp n^{-m/2}\tau^{mn}, \qquad \left(\|B_{m,n}\|_2^2\right)'' \le -cn\|B_{m,n}(u_m)\|_2^2.$$

The first derivative of $B_{m,n}$ itself is zero at the barycentre by exchange symmetry. Therefore

$$\langle B''_{m,n}, B_{m,n}\rangle \le -cn\|B_{m,n}(u_m)\|_2^2. \tag{55}$$

We now pass from pure oscillator inputs to admissible relative perturbations. At each fixed even Gaussian Hermite degree $2j$, the pairing of H_n with e_{2j} is $O_j(n^{C_j}\lambda^n)$. Indeed its generating function in n is a fixed polynomial in the generating variable times $\exp(r_0 z^2/2)$, by integrating the Gaussian against the fixed polynomial e_{2j}. Extracting the even coefficient and multiplying by $\sqrt{n!}$ gives the bound using $r_0 = \lambda^2$. To cancel all the prescribed moments, subtract a linear combination of a fixed finite set of $\cos(\omega_j x)$ with distinct nonnegative ω_j^2. Their matrix of low even Hermite coefficients is invertible, since these coefficients are proportional, row by row and column by column, to a Vandermonde matrix in ω_j^2. Thus all correction coefficients are $O(n^C\lambda^n)$. Call the corrected oscillator $\widetilde{H}_n$. Set $g_n = e^{-\sqrt{n}}\widetilde{H}_n$. Oscillator recurrences and the one-dimensional Sobolev supremum bound show that the derivatives of each fixed order of H_n grow at most polynomially in n. Hence $g_n \to 0$ in each C_b^k, and $-\log((1+g_n)\gamma)$ has second derivative bounded below by a fixed positive constant for all large n. The correction enforces mass and variance as well as the specified higher moments.

To compare all output terms at u_m, use the conditional subset formula and its differentiated version. For an uncorrected size-j subset with input H_n, its norm and those of its first two weight derivatives are at most a polynomial in n times $\tau_a(ja)^{jn/2}$. For the derivatives, use the degree factors of order at most two, paid by the tensor total-number norms. The single-factor estimates follow from the oscillator Gram formula and tilting the level by $\exp(c/(n+1))$ as in Section 3. At $j = 1, b = a$ one can use the coarser bound $S_n(b) \le C\max(\nu,\tau)^n$, which is uniform as $b \downarrow a$ and has the same exponential rate. For the corrected inputs, any term involving a correction has norm and first two derivative norms $O(n^C\lambda^n)$ before the scalar amplitude, using the bounded conditional Gaussian kernel and its derivatives at fixed $q < 1$. By the strict subset comparison following (41), all $j < m$ terms, as well as the corrections, are exponentially smaller than the full m-factor term, even after accounting for the factors $e^{-j\sqrt{n}}$ and polynomial factors in n.

Consequently the leading part of the entropy pairing $\langle R'', R\rangle$ is $e^{-2m\sqrt{n}}$ times (55), with an error smaller than its magnitude. At equal weights $R' = 0$ by exchange symmetry. The other entropy error satisfies

$$|\langle R'', \psi(R)\rangle| \le C\|R\|_\infty \|R''\|_2 \|R\|_2.$$

The two L^2 norms are bounded by a polynomial in n times $e^{-m\sqrt{n}}\tau^{mn/2}$, and $\|R\|_\infty \le e^{-\sqrt{n}}\,\mathrm{poly}(n)$, so this error is negligible compared with $ne^{-2m\sqrt{n}}\|B_{m,n}(u_m)\|_2^2$. It follows that $F''_{q,g_n}(u_m) > 0$. The estimates are uniform on the unit tangent sphere, or one may use that a permutation-invariant Hessian at the barycentre is scalar on the tangent space. This completes Theorem 9.2.

10 Critical concavity neighborhoods by copy count

10.1 Statement and tools

Let γ be the standard Gaussian law on the line. An admissible relative perturbation g is real, even and bounded, with $\int g\,d\gamma = \int x^2 g\,d\gamma = 0$. For $\|g\|_\infty < 1$, independent $X_i \sim (1+g)\gamma$, and an independent $G \sim \gamma$, set

$$F_{q,g}(w) = h\left(\sqrt{q}\sum_{i=1}^{m}\sqrt{w_i}X_i + \sqrt{1-q}G\right), \qquad w_i > 0, \quad \sum_i w_i = 1, \quad 0 < q < 1.$$

Here h is differential entropy and $u = (1/m, \dots, 1/m)$. Put

$$\lambda_a = \frac{a}{1+\sqrt{1-a^2}}, \qquad \tau_a(b) = \frac{\sqrt{1-a^2}-1+b}{\sqrt{1-a^2}+1-b}, \qquad \tau_{q_m/m}(q_m)^m = \lambda_{q_m/m}.$$

The last equation defines a unique $q_m \in (0,1)$ Sections 9.1–9.5. In particular $q_3 = 0.65767078\dots$ and $q_4 = 0.71210929\dots$.

Theorem 10.1. *For $m = 3, 4$ there are $\delta, \rho > 0$ such that every admissible g with $0 < \|g\|_\infty \le \delta$ has strictly negative tangent Hessian for all $0 < q \le q_m$ and $\|w - u\|_2 \le \rho$. For every $m \ge 5$ at $q = q_m$, each neighborhood of u contains a weight w and a fixed nonzero tangent direction at which F_{q,g_n} has strictly positive second derivative for a sequence of smooth strongly log-concave densities $(1+g_n)\gamma$. The g_n converge to zero in each fixed C_b^k norm and may match any prescribed finite Gaussian moments.*

The theorem of Section 9.1 gives a common full weight ball below q_m for every $m \ge 3$, uniform radial concavity at q_m, and positive exact barycentric curvature above q_m. Thus the question at q_m concerns tangent directions at nearby weights. Sections 6 and 7 and the Evidence Press candidate [5, Section 9.3] discuss the distinction between barycentric and near-face behavior. We use the oscillator estimates and Gaussian product coefficients as developed in [7, 5] and Sections 9.1–9.5; the proof below gives the additional estimates needed at the critical signal.

Write T_ρ for the Ornstein–Uhlenbeck operator and $\mathcal{N}$ for its number operator, with multipliers ρ^n and n on the normalized Hermites e_n. Every unmarked norm is for γ. The output relative density is $1 + R$, with $R = L + M$ and $L = \sum_i T_{\sqrt{qw_i}}g$. The subset H of size j contributes, at Hermite degree $n = \sum_{i\in H} n_i$, the coefficient summands

$$\sqrt{\frac{n!}{\prod_{i\in H} n_i!}}\prod_{i\in H}(qw_i)^{n_i/2}\widehat{g}(n_i). \tag{56}$$

The sum M contains $j \ge 2$. These formulas follow from Gaussian conditioning. For any levels $b_i \in (0,1)$ such that $\sum_{i\in H} qw_i/b_i \le 1$, the subset is the Gaussian conditional projection of a tensor of the independent $T_{\sqrt{b_i}}g$ factors, with coefficients $\sqrt{qw_i/b_i}$ in the projection and an independent Gaussian to fill any unused variance. The projection is an L^2 contraction and commutes with tensor total number. In a second weight derivative, with the b_i frozen at the evaluation point, the summands in (56) acquire a multiplier bounded by $C(\sum n_i)^2$ on a fixed positive weight neighborhood. Thus a second derivative after division by the output number costs at most one total number on the tensor. This observation applies also after adding or removing a fixed Gaussian attenuation in the output, by changing the levels qw_i in the estimate.

For a unit tangent v, primes denote weight derivatives and put $E = \langle L'', L\rangle$. The linear scalar multipliers at degree $2k \ge 4$ are

$$q^k \sum_i w_i^k, \qquad q^k k(k-1) \sum_i v_i^2 w_i^{k-2}$$

for L and L'', respectively. On compact positive weights $\|\mathcal{N}^{-1} L''\| \le C E^{1/2}$. All inverse number operators act on positive degrees. The entropy derivative is

$$F'' = -\langle R'', R + \psi(R)\rangle - \int \frac{(R')^2}{1+R}\, d\gamma, \qquad \psi(z) = \log(1+z) - z. \tag{57}$$

Gaussian conditioning also gives spatial derivative bounds $\|R\|_\infty + \|R_x\|_\infty + \|R_{xx}\|_\infty \le C\|g\|_\infty$ when q stays below one. These bounds hold for L and M as well. They follow by differentiating the Gaussian conditional density in its mean; the conditional covariance is bounded below by $1-q$. Weight derivatives on a compact interior set are justified by the same Gaussian formula, with polynomial growth in the output variable.

Here are the oscillator bounds from Sections 9.3 and 9.4 used below. With reference level a and $D = \|T_{\sqrt{a}} g\|_2$, if $b > a$ and $p < \log \tau_a(b) / \log \lambda_a$, $p > 0$, then

$$\|T_{\sqrt{b}} g\|_2 \le C \delta^{1-p} D^p, \qquad \delta = \|g\|_\infty. \tag{58}$$

There is also an L^P bound of this form, for $P > 2$ close to two and $b > a$ close to a, with $pP/2 > 1$. These strict margins are uniform on compact level sets. In the weighted endpoint estimate, set $X = \|(\mathcal{N}+1) T_{\sqrt{a}} g\|_2$. One splitting integer K, determined only by $a, X/\delta$, gives

$$\|\mathcal{N}^j T_{\sqrt{b}} g\|_2 \le C\delta (K+1)^j \tau_a(b)^{K/2}, \quad j = 0, 1, \qquad \delta (K+1) \lambda_a^{K/2} \le C X.$$

The constants are uniform for levels with $b - a > 0$ bounded away from zero and $b < 1$ bounded away from one. It follows from the same integer K that for unequal levels and each $i \in H$,

$$\|\mathcal{N} T_{\sqrt{b_i}} g\|_2 \prod_{\substack{k \in H \\ k \ne i}} \|T_{\sqrt{b_k}} g\|_2 \le C \delta^{|H|-1} X \quad \text{if} \quad \prod_{k \in H} \tau_a(b_k) \le \lambda_a. \tag{59}$$

10.2 The endpoint for three and four copies

We work for q in a short compact interval ending at q_m; other signals are covered by Sections 9.1–9.5. Write $a_0 = q/m$ and order the weights $w_{(1)} \le \cdots \le w_{(m)}$. For every unit tangent at least $1/m$ of its squared norm lies on the weights indexed $2, \dots, m$, since the remaining component is minus the sum of the others. By the two linear multipliers above,

$$E \ge c X^2, \qquad X = \|(\mathcal{N}+1) T_{\sqrt{a}} g\|_2, \qquad a = \min\{a_0, q\sqrt{w_{(2)} w_{(m)}}\}. \tag{60}$$

The absence of degrees below four justifies $\mathcal{N} + 1$ in this inequality.

For $m = 3$, the comparison level equals a_0 near u. If $w_{(2)} = 1/3 - e \le 1/3$, then $w_{(3)} \ge 1/3 + 2e$ and their product is at least $1/9$ for e small. For $w_{(2)} \ge 1/3$ the same product bound is immediate. At the full subset use all $b_i = q$. Then $\tau_a(q)^3 \le \lambda_a$ by $q \le q_3$.

For $m = 4$, put $z = w - u$. The ordering gives $w_{(2)} + w_{(4)} \ge 1/2$. Using $z_{(2)} z_{(4)} \ge -\|z\|_2^2/2$ yields

$$a \ge a_0 (1 - 4\|z\|_2^2 + O(\|z\|_2^4)). \tag{61}$$

For the full subset instead take

$$S = \sum_i \sqrt{w_i}, \qquad b_i = q\, S\sqrt{w_i}.$$

The conditional projection criterion holds with equality. All levels tend to q at u. Since $\log \tau_a(b)$ is concave in $b < 1$ for $b \geq a$,

$$\sum_i \log \tau_a(b_i) \leq 4 \log \tau_a(qS^2/4).$$

Here $S^2/4 = 1 - \|z\|_2^2 + O(\|z\|_2^3)$. The decrease from q to $qS^2/4$ outweighs the loss in the comparison level a in (61). For the explicit inequality, evaluate derivatives at $a_0 = q/4, b = q$, and write $d = \sqrt{1 - a_0^2}$ and $D_0 = d^2 - (1-q)^2$. It is enough that

$$4q\, \partial_b \log \tau_{a_0}(q) > 4a_0\, \partial_a (\log \lambda_a - 4 \log \tau_a(q))\big|_{a=a_0}.$$

The derivatives obey $a_0 \partial_a \log \lambda_{a_0} = 1/d$, $-a_0 \partial_a \log \tau_{a_0}(q) = 2a_0^2(1-q)/(dD_0)$, and $q\partial_b \log \tau_{a_0}(q) = 2qd/D_0$. The desired difference has the sign of

$$2qd^2 - D_0 - 8a_0^2(1-q) = \frac{q^2(9+6q)}{16} > 0.$$

Thus in a common small ball $\prod_i \tau_a(b_i) \leq \lambda_a$, because the same inequality holds at u for $q \leq q_4$. All Taylor estimates are uniform for q near q_4.

For either copy count the proper subsets have strict margin with the common input level $b_H = q \sum_{i \in H} w_i$. Indeed at the centre their comparison is strictly below the full critical value, as follows from strict convexity of $j \log \tau_{a_0}(j a_0)$ from $j = 1$ to m Sections 9.1–9.5. By (59), the bounds just proved for the full and proper subsets give

$$\|\mathcal{N} M\|_2 + \|\mathcal{N}^{-1} M''\|_2 \leq C \delta X. \tag{62}$$

Use the conditional contraction with one tensor number and freeze the b_i when applying the derivative factors in (56). In particular $|\langle M'', M \rangle| \leq C\delta^2 X^2$.

We next show that the two linear–mixed pairings are $O(\delta E)$ with a strict smoothing transfer. For a fixed $p_0 < 1$, apply $T_{\sqrt{p_0}}$ to the mixed factor and $T_{1/\sqrt{p_0}}$ to the linear factor. The latter is the inverse coefficient multiplier on its smoothed inputs and is well defined for p_0 close to one. At the barycentre the available exponents in (58), relative to a_0, for size j and the linear factor are respectively

$$j \frac{\log \tau_{a_0}(p_0 j a_0)}{\log \lambda_{a_0}}, \qquad \frac{\log \tau_{a_0}(a_0/p_0)}{\log \lambda_{a_0}}. \tag{63}$$

For the full subset at $q = q_m, p_0 = 1$ their sum is two. The numerator of the sum has strictly positive derivative in $\log p_0$ there, because

$$mq\, \partial_b \log \tau_{a_0}(q) > a_0\, \partial_b \log \tau_{a_0}(a_0).$$

For instance $b\partial_b \log \tau_a(b)$ is increasing for $b \geq a$. The denominator is negative, so decreasing p_0 gives a strict sum greater than two. The same holds for each proper subset, already strict at $p_0 = 1$. Choose one such p_0, then a fixed reference $a_{\mathrm{ref}} = q(1/m - \xi)$ slightly below a_0. The margins persist by continuity near the critical signal and the centre. In particular $a_{\mathrm{ref}} \leq a$. Spend part of the margins on number powers incurred by the weight derivatives. Thus there are exponents $\alpha, \beta > 0$

with $\alpha+\beta>2$, $\beta<1$, and $\alpha<2$, so the transferred size-j norms, including their needed derivatives, are respectively at most

$$C\delta^{j-\alpha}D^{\alpha}, \qquad C\delta^{1-\beta}D^{\beta}, \qquad D=\|T_{\sqrt{a_{\mathrm{ref}}}}g\|_2.$$

Indeed at p_0 sufficiently close to one the first exponent for the full subset is slightly above one, the proper subset exponents have slack, and the second exponent is slightly below one. Because $D\le\delta\le 1$, the product is at most $C\delta D^2$. This applies to both pairings.

Finally the L^P oscillator bound with that same reference level, chosen sufficiently close to a_0, gives $\|\mathcal{N}\psi(L)\|_2\le C\delta D$. For completeness take $P>2$ and $pP/2>1$ from the strict L^P margin. Factor each linear term through a slightly larger level to pay its first two spatial derivatives, then interpolate these norms between L^P and the $O(\delta)$ sup bounds. Products of two such norms at $U>4$ with $2Pp/U>1$ give $C\delta D$ in L^2. This is valid also with a factor of x, using a finite Gaussian L^A norm with $1/A+2/U=1/2$. Apply this to $\mathcal{N}\psi(L)=\psi'(L)(-L_{xx}+xL_x)-\psi''(L)L_x^2$.

Put $W=\psi(L+M)-\psi(L)$. The number chain rule and the spatial sup bounds give

$$\|\mathcal{N}W\|_2\le C\delta\big(\|\mathcal{N}M\|_2+\|(1+|x|)M\|_2+\|M_x\|_2\big)\le C\delta\|(\mathcal{N}+1)M\|_2.$$

In the chain-rule difference, the terms with $\mathcal{N}M$ and $R_x^2-L_x^2$ use $|\psi'(R)|\le C\delta$ and $|R_x+L_x|\le C\delta$. The other terms contain M times $\mathcal{N}L=O(\delta(1+|x|))$ or $L_x^2=O(\delta^2)$. The last inequality follows from the Hermite recurrence for multiplication by x and for the spatial derivative. Now (62) controls this difference by $C\delta^2X$. Pairing the second derivatives through $\mathcal{N}^{-1}$ proves that the logarithmic error in (57) is $O(\delta E)$. Thus the restoring term E dominates all errors for a common small δ, which proves the positive part of Theorem 10.1.

10.3 The obstruction for five or more copies

Fix $m\ge 5$, $q=q_m$, $a=q/m$, and abbreviate $\lambda=\lambda_a$, $\tau=\tau_a(q)$, so that $\tau^m=\lambda$. Use the Lebesgue oscillator modes H_n at reference level a, with $n\to\infty$ even. For this basis put

$$\nu_a(b)=\frac{1+b-\sqrt{1-a^2}}{1+b+\sqrt{1-a^2}}, \qquad S_n(b)=\|T_{\sqrt{b}}H_n\|_2^2.$$

The oscillator diagonal generating function of Sections 9.1–9.5 gives $S_n(b)\le C\max\{\nu_a(b),|\tau_a(b)|\}^n$. The maximum is $\nu_a(b)$ for $0<b\le a$ and $\tau_a(b)$ for $b\ge a$. Fixed number powers cost only polynomial factors in n after enlarging b by a factor $\exp(c/(n+1))$; on compact level sets this enlargement changes the geometric bound by a constant factor.

For a small fixed $e>0$, take

$$w=(1/m-e,1/m-e,1/m+2e/(m-2),\dots,1/m+2e/(m-2)), \qquad v=(1,-1,0,\dots,0)/\sqrt{2}. \tag{64}$$

Put $a_-=q(1/m-e)$ and $a_+=q(1/m+2e/(m-2))$. Since $m\ge 5$, $2/(m-2)<1$, so

$$b_\times:=\sqrt{a_-a_+}=a-c_me+O(e^2) \quad \text{with } c_m>0.$$

The linear second derivative along v differentiates only the two small weights. For the oscillator input, each pairing of that derivative with a linear factor at level a_i is a number-weighted squared norm at level $\sqrt{a_-a_i}\le b_\times$. Thus, with polynomial factors allowed in n, its magnitude is bounded by $\nu_a(b_\times)^n$. In particular

$$\log\nu_a(b_\times)\le\log\lambda-ce \tag{65}$$

for all sufficiently small $e > 0$. The norm squared of the linear second derivative alone has at most this exponential rate as well.

Let $B_{m,n}(w)$ be the full conditional product with each original input H_n. Section 9.5 expresses its norm squared as a positive Gaussian quadratic coefficient, and gives the following C^2 coefficient asymptotic near u:

$$\|B_{m,n}(w)\|_2^2 = n^{-m/2}\tau^{mn}\exp(n(J(w)-m))\,(A(w)+o_{C^2}(1)), \tag{66}$$

with smooth $A > 0$ and $J(u) = m$, $J'(u) = 0$ on tangent directions, and strictly negative tangent Hessian of J at u. These coefficient calculations apply unchanged at $q = q_m$: they use strict positivity of an oscillator Gram matrix near u and $q \in (0,1)$, before any comparison of the full product with other subsets. In particular $J(w) - m = O(e^2)$. At the weights (64), the exchange of the first two coordinates gives $J'(w)[v] = B'_{m,n}(w)[v] = 0$. By continuity $J''(w)[v,v] < 0$. Thus (66) gives

$$\langle B''_{m,n}, B_{m,n}\rangle = \tfrac{1}{2}(\|B_{m,n}\|_2^2)'' \le -cn\|B_{m,n}(w)\|_2^2. \tag{67}$$

We also need the derivative norm at these skew weights to retain the $J(w)$ exponent. For two sets of weights w, w', the Gram pairing $\langle B_{m,n}(w), B_{m,n}(w')\rangle$ has the same Gaussian coefficient representation as in Sections 9.1–9.5, with cross covariance $q\sqrt{w}(\sqrt{w'})^\top$. Each block vector is unit, so the Gaussian determinant prefactor is independent of both weights. After multiplying the first block variables by i and the second by $-i$, the coefficient is unchanged. The transformed entries are strictly positive on a fixed small neighborhood of (u,u), by the strict positivity at equality established in Sections 9.1–9.5. Expand the coefficient into edge and loop monomials in these positive entries. In every monomial the sum of entry exponents is mn, since the total variable degree is $2mn$. Each derivative of order at most four of such a monomial is therefore bounded, in absolute value, by a fixed polynomial in n times the monomial itself. Here the entries are smooth and bounded away from zero on a compact subneighborhood. Positivity permits summing this bound. Differentiating twice in each weight block and then putting $w' = w$ proves

$$\|B''_{m,n}(w)\|_2 \le Cn^2\|B_{m,n}(w)\|_2. \tag{68}$$

The equality between the differentiated Gram pairing and the derivative norm follows from smooth conditional Gaussian integrals at these positive weights.

We now compare all other terms with the full pairing (67). At u, a proper mixed subset has squared norm rate at most $\tau_a(ja)^j < \lambda$ for $2 \le j < m$. The same strict exponential margin holds after moving the weights slightly, and after taking two weight derivatives: apply the conditional contraction and the polynomial number cost to the oscillator bounds above. Since the other subset norms are at most $\lambda^{n/2}\exp(O(e)n)$ times fixed polynomial factors, any pairing involving a proper mixed subset is negligible relative to (67) at sufficiently small fixed e.

For the full–linear pairings one has a strict exponential gap by transferring $T_{\sqrt{p_0}}$ to the full factor and $T_{1/\sqrt{p_0}}$ to the linear one, as in the positive proof. In fact the bound at u before polynomial factors is

$$[\tau_a(p_0 q)^m\,\tau_a(a/p_0)]^{n/2}\,.$$

For $p_0 < 1$ close to one the bracket is strictly below λ^2 by the derivative inequality used in (63). This remains strict in a small weight neighborhood, including the number factors in two derivatives, which cost polynomial factors in n. Thus both full–linear pairings are negligible on the full exponential scale at (64). The pure linear pairing is negligible by (65), since its loss is linear in e and $J(w) - m = O(e^2)$. The same argument bounds the norm of the linear second derivative by an exponentially smaller quantity than $\|B_{m,n}(w)\|_2$.

As in Sections 9.1–9.5, subtract from H_n a linear combination of finitely many fixed bounded cosine modes to cancel any prescribed even Gaussian moments, including mass and variance. The coefficient of each correction is $O(n^C\lambda^n)$: the low-degree Gaussian pairings of H_n have this order by their one-dimensional oscillator generating function, and the moment matrix for distinct squared cosine frequencies is Vandermonde up to fixed nonzero factors. After the correction set $g_n = e^{-\sqrt{n}}\widetilde{H}_n$. The fixed-order oscillator derivative norms grow polynomially in n; thus $g_n \to 0$ in every fixed C_b^k and $(1+g_n)\gamma$ is strongly log-concave for all large n. In each subset, terms containing a correction and their first two weight derivatives have L^2 norms at most $O(n^C\lambda^n)$ before the scalar amplitude. This follows by bounded Gaussian conditioning of the inputs and differentiation of the conditional density. They have a strict exponential gap from the full norm in (66) for small e.

The scalar powers $e^{-j\sqrt{n}}$ are subexponential in n and do not affect the strict gaps just proved. Consequently for $R = L + M$ built from g_n, the leading pairing $\langle R'', R\rangle$ is $e^{-2m\sqrt{n}}\langle B''_{m,n}, B_{m,n}\rangle$, up to a smaller error. By (68) and the comparison for L'',

$$\|R''\|_2 \le n^C e^{-m\sqrt{n}}\|B_{m,n}(w)\|_2, \qquad \|M\|_2 \le Ce^{-m\sqrt{n}}\|B_{m,n}(w)\|_2.$$

It remains to bound the nonlinear logarithm without using the possibly larger L^2 norm of L. The oscillator L^P mode estimate in Sections 9.1–9.5 gives, at levels b sufficiently close to a, $\|T_{\sqrt{b}}H_n\|_P \le C\zeta^{n/2}$ with $\zeta^{P/2} < \lambda$. Indeed ζ can be chosen above the larger of the two Mehler radii in that estimate, and the strict L^P margin gives the displayed inequality for $P > 2$ sufficiently close to two. For levels below a one may first enlarge the level, then apply contraction in L^P. Interpolating to L^4 using the polynomial oscillator sup norm shows that $\|L\|_4^2$ has an exponential gap below $\lambda^{n/2}$, even allowing the changes in weights and cosine corrections. Since $J(w) - m = O(e^2)$, this is exponentially smaller than $e^{-m\sqrt{n}}\|B_{m,n}(w)\|_2$ at small fixed e.

On the other hand the mean value inequality for ψ gives

$$\|\psi(R)\|_2 \le C\|L\|_4^2 + C(\|L\|_\infty + \|M\|_\infty)\|M\|_2.$$

The displayed sup norms are at most $e^{-\sqrt{n}}$ times a polynomial in n by Gaussian conditioning. Together with the bound on R'', this proves that $\langle R'', \psi(R)\rangle$ is negligible relative to the contribution of (67). Finally, $R'(w)[v] = 0$ by the exchange symmetry at (64). Equation (57) therefore gives $F''_{q,g_n}(w)[v,v] > 0$ for all sufficiently large n. Since $e > 0$ can be chosen arbitrarily small, this proves the negative part of Theorem 10.1.

11 Critical thresholds in higher dimensions

11.1 Results

Let γ_d be standard Gaussian measure on $\mathbb{R}^d$. Write $\mathcal{T}_b = T_{\sqrt{b}}$ for the Ornstein–Uhlenbeck operator with multiplier $b^{n/2}$ on Gaussian Hermite degree n, and write $\mathcal{N}$ for the number operator. All unmarked norms are for γ_d. For $0 < a < b < 1$, set

$$d_a = \sqrt{1-a^2}, \quad B_a = d_a^{-1}, \quad \lambda_a = \frac{a}{1+d_a}, \quad \tau_a(b) = \frac{d_a - 1 + b}{d_a + 1 - b}, \quad \nu_a(b) = \frac{1 + b - d_a}{1 + b + d_a}.$$

Define $\theta = \log \tau_a(b)/\log \lambda_a \in (0,1)$. For nonzero bounded g, write $Y = \|g\|_\infty$ and $X = \|\mathcal{T}_a g\|_2$.

Theorem 11.1. *For every fixed d, a, b as above,*

$$\|\mathcal{T}_b g\|_2^2 \le CY^{2-2\theta}X^{2\theta}(1 + \log(Y/X))^{[d(1-\theta)-1]/2}.$$

The exponent of the logarithmic factor is sharp. For an integer $m \geq 2$ with $\tau_a(b)^m = \lambda_a$, the corresponding inequality with one output number factor is

$$\|\mathcal{N}\mathcal{T}_b g\|_2 \|\mathcal{T}_b g\|_2^{m-1} \leq C Y^{m-1} X_1 \big(1 + \log_+(Y/X_1)\big)^{[d(m-1)-m]/4}, \qquad X_1 = \|(\mathcal{N}+1)\mathcal{T}_a g\|_2.$$

Its logarithmic exponent is also sharp.

For $m = 2$ the logarithmic gain of exponent $-1/4$ in dimension one is optimal. The unweighted linear bound at the critical relation holds in dimensions one and two, and fails for every $d \geq 3$. In dimensions $d = 1, 2$ the best critical linear constant has order $a^{-d/8}$ as $a \downarrow 0$. The corresponding entropy conclusion uses a stronger collective estimate and has no such dimension restriction.

A bounded perturbation g is admissible if it is centrally even and $(1+g)\gamma_d$ has mass one and covariance I_d. For $\|g\|_\infty < 1$ and independent variables $X_i \sim (1+g)\gamma_d$, $G \sim \gamma_d$, set

$$F_{q,g}(w) = h\left(\sqrt{q}\sum_{i=1}^{m}\sqrt{w_i}X_i + \sqrt{1-q}G\right).$$

Here h is differential entropy, $w_i \geq 0, \sum_i w_i = 1$, and $0 < q < 1$. The barycentre is $u = (1/m, \ldots, 1/m)$. Define q_m by $\tau_{q_m/m}(q_m)^m = \lambda_{q_m/m}$, with $q_2 = 4/7$.

Theorem 11.2. *Fix $d \geq 1$. For $m = 2$, one bounded Gaussian-relative neighborhood has strictly negative weight curvature for non-Gaussian admissible inputs at all mixing weights and all $0 < q \leq 4/7$, including the continuous one-sided endpoint derivatives in the weight. Each $q > 4/7$ admits counterexamples arbitrarily close in relative density.*

For each $m \geq 3$ and every $Q < q_m$, a common bounded Gaussian-relative neighborhood has negative definite tangent Hessian on a common open weight ball about u, uniformly for $0 < q \leq Q$. At $q = q_m$ there is uniform radial concavity on such a ball, with negative definite Hessian at u. A common full weight ball through q_m exists for $m = 3, 4$ and fails for $m \geq 5$. At each fixed $q > q_m$, arbitrarily small smooth strongly log-concave relative perturbations give positive definite Hessian at u. Both types of obstructions may match any fixed finite Gaussian moments.

Luo [7] introduced the oscillator comparison underlying the two-copy boundary. The Evidence Press manuscript [5, Remark 5.4 and Section 9.3] gives the one-dimensional bounded-function estimates and asks for the higher-dimensional entropy endpoint. Its proposed scalar extension has a logarithmic loss. Theorem 11.1 gives its sharp size, including the range where the loss is necessary. The scalar loss would prevent direct use of the product of separate norms at the entropy boundary. Instead we estimate the conditional product directly. The weight results in Sections 6, 9.1–9.5 and 10.1–10.3 supply the one-dimensional comparisons and the scalar signal algebra.

11.2 Oscillator shells and the scalar estimate

Use the tensor Lebesgue oscillator basis at scale B_a, starting with

$$H_k(x) = c_a e^{-B_a x^2/2} \mathrm{He}_k(\sqrt{2B_a}\, x)/\sqrt{k!}$$

in dimension one, with c_a chosen to make it orthonormal. Write P_k for the projection onto oscillator degree k in dimension d, and $\mathcal{H}P_k = kP_k$. These degrees differ from the Gaussian Hermite degrees of $\mathcal{N}$.

The Gaussian covariance form $\langle \mathcal{T}_b H_j, \mathcal{T}_b H_k\rangle$ has the generating function, with coefficients divided by $\sqrt{j!k!}$, proportional to

$$\exp\left(\frac{\nu_a(b)(z+w)^2 - \tau_a(b)(z-w)^2}{4}\right).$$

This follows from a two-dimensional Gaussian integral and tensors over coordinates. At $b = a$ this is diagonal with weights proportional to λ_a^k. For $b > a$, scaling both generating variables by $\tau_a(b)^{-1/2}$ identifies the remaining form with Gaussian multiplication on $L^2(dx)$. In one coordinate the multiplication is e^{-cx^2}, where

$$\frac{B_a - c}{B_a + c} = \frac{\nu_a(b)}{\tau_a(b)}.$$

Indeed the generating calculation for this multiplier replaces $\nu_a(b)/\tau_a(b)$ by the left side and keeps the coefficient of $-(z-w)^2/4$ equal to one. Thus, up to a positive constant,

$$\|\mathcal{T}_b f\|_2^2 = C\|e^{-c|x|^2/2}\tau_a(b)^{\mathcal{H}/2} f\|_{L^2(dx)}^2. \tag{69}$$

For a one-dimensional mode the diagonal generating function is proportional to $((1-z\nu_a(b))(1-z\tau_a(b)))^{-1/2}$. This can be obtained by substituting a standard circular Gaussian and its conjugate for z, w and averaging. Since $0 < \nu_a(b) < \tau_a(b)$, its kth coefficient is at most $C(k+1)^{-1/2}\tau_a(b)^k$. Equivalently $\|e^{-cx^2/2}H_k\|_{L^2(dx)}^2 \le C(k+1)^{-1/2}$. It follows that on a d-dimensional oscillator shell

$$\|e^{-c|x|^2/2}P_k\|_{L^2(dx)\to L^2(dx)} \le C(k+1)^{-1/4}. \tag{70}$$

To see this, partition the multi-indices of total k into d groups by choosing a coordinate of degree at least k/d. In a group discard the Gaussian weights in all other coordinates and use orthogonality there. The degrees in those coordinates fix the remaining one on the shell. The one-dimensional bound applies to that degree, and summing the finitely many groups gives (70). The bound holds for any fixed $c > 0$, since the multiplication generating calculation and the diagonal estimate also apply directly to e^{-cx^2} with parameter $\nu/\tau = (B_a - c)/(B_a + c) \in (-1, 1)$.

For a unit $L^2(dx)$ function f on shell k, the oscillator ladder relations give $\int |x|^{2j}|f(x)|^2 dx \le C_j(k+1)^j$ for each fixed integer $j \ge 0$. Weighted Cauchy–Schwarz using $(1+|x|^2/(k+1))^j$, $j > d/2$, gives $\|f\|_{L^1(dx)} \le C(k+1)^{d/4}$. By duality and by orthogonality at a,

$$\|P_k g\|_{L^2(dx)} \le C\min\{Y(k+1)^{d/4},\, X\lambda_a^{-k/2}\}. \tag{71}$$

The projections can be defined by pairing with integrable oscillator functions. All expansions follow first for bounded compactly supported inputs and then by cutting off g, using $L^2(\gamma_d)$ convergence and the absolutely summable estimates.

We also need a weighted variant:

$$\|P_k g\|_{L^2(dx)} \le C\min\{Y(k+1)^{d/4}, X_1(k+1)^{-1}\lambda_a^{-k/2}\}. \tag{72}$$

For levels $b_0 < a$ close to a, scale the Gram by $\nu_a(b_0)^{-1/2}$ and make the phase substitution $z \mapsto iz, w \mapsto -iw$. This is a unitary oscillator rotation and swaps the two normalized parameters in the generating function. The resulting normalized form is again Gaussian multiplication, since $|\tau_a(b_0)| < \nu_a(b_0)$. Hence $\|\mathcal{T}_{av}P_k\|^2 \le C\lambda_a^k e^{-ck(1-v)}$ for v near one, using the positive derivative of ν_a. For smaller v use monotonicity of the OU norm. The identity

$$\|(\mathcal{N}+1)^{-1}\mathcal{T}_a f\|_2^2 = \int_0^1 (-\log v)\|\mathcal{T}_{av} f\|_2^2\, dv$$

then bounds this operator on the shell by $C(k+1)^{-1}\lambda_a^{k/2}$. Pair with $(\mathcal{N}+1)\mathcal{T}_a g$ and use the diagonal form at a to obtain (72). Likewise a small enlargement $b\exp(c_0/(k+1))$ in (69) absorbs a number power:

$$\|\mathcal{N}^j\mathcal{T}_b P_k\|_{L^2(dx)\to L^2(\gamma_d)} \le C_j(k+1)^{j-1/4}\tau_a(b)^{k/2}. \tag{73}$$

Indeed $n^{2j} \le C_j(k+1)^{2j}\exp(c_0 n/(k+1))$ at Gaussian degree n, and the enlargement changes $\tau_a(b)^k$ by a bounded factor.

Split the unweighted shell sum at a K where $Y(K+1)^{d/4}\lambda_a^{K/2}$ is comparable to X. Use $K = O(1)$ if X/Y is bounded below. The two exponential tails from (71), (73) give

$$\|\mathcal{T}_b g\|_2 \le CY(K+1)^{(d-1)/4}\tau_a(b)^{K/2}.$$

Here the lower-shell ratio is geometric since $\tau_a(b) > \lambda_a$, and the upper-shell ratio is geometric since $\tau_a(b) < 1$. Substitution of the crossover and $K+1 \asymp 1+\log(Y/X)$ gives the first inequality. For (72) the crossover is $Y(K+1)^{1+d/4}\lambda_a^{K/2} \asymp X_1$. At that crossover $\|\mathcal{N}^j\mathcal{T}_b g\| \le CY(K+1)^{j+(d-1)/4}\tau_a(b)^{K/2}$. Taking the indicated product at $\tau_a(b)^m = \lambda_a$ gives the second inequality of Theorem 11.1.

11.2.1 Sharpness from bounded radial examples

Let $\alpha = d/2-1$, $l = \alpha+1$, and $u = B_a|x|^2$. Define

$$c_j = [z^j](1-z)^l(1+z), \qquad G_n(x) = e^{-u/2}\sum_{j\ge 0} c_j L_{n+j}^{(\alpha)}(u).$$

If l is an integer the sum is finite. Otherwise $c_j = O((j+1)^{-l-1})$ by the binomial coefficients. The radial Laguerre function of index $n+j$ lies in oscillator degree $2(n+j)$, by its differential equation, and its $L^2(dx)$ norm is a fixed constant times $(\Gamma(n+j+\alpha+1)/(n+j)!)^{1/2}$. Consequently the sum converges absolutely in $L^2(dx)$.

There is a bounded representative with

$$\|G_n\|_\infty \le C(n+1)^{-1/2}. \tag{74}$$

Using the Laguerre generating function [12], take the nth angular Fourier coefficient of

$$(1-\overline{z})^l(1+\overline{z})(1-z)^{-l}\exp\left(-\frac{u(1+z)}{2(1-z)}\right) \qquad (z=\rho e^{it}).$$

This coefficient is the defining sum with extra factors ρ^{n+2j}. Its limit in $L^2(dx)$ follows from the summability just proved. Pointwise the integral converges by bounded convergence to a constant phase times

$$\frac{1}{2\pi}\int_0^{2\pi}(1+e^{-it})\exp\left(-i(n+l)t - \frac{iu}{2}\cot(t/2)\right)dt.$$

Here the fractional powers are the analytic powers in the disk, and their boundary phase ratio is $e^{-il(t-\pi)}$ for $0 < t < 2\pi$. The integrand at $\rho < 1$ is bounded by two, since the power ratio has unit modulus and the exponential has modulus at most one.

For the uniform estimate put $N_0 = n+l$. The phase derivative is $-N_0 + u/(4\sin^2(t/2))$. If $0 < u \le N_0$, split at $\sin^2(t/2) = u/(2N_0)$. On the inner end intervals the second derivative has absolute value at least cN_0. On the rest the first derivative has absolute value at least $N_0/2$, and is monotone on each half. If $u \ge N_0$, split into dyadic rings about $t = \pi$. On a ring of distance comparable to $r \le 1$, the second derivative has magnitude at least $cN_0 r$, while the amplitude and its total variation on the ring are $O(r)$. The second-derivative estimate gives $CN_0^{-1/2}r^{1/2}$; summing the rings gives $CN_0^{-1/2}$. The remaining intervals use the second-derivative estimate with r of order one. These are the elementary van der Corput estimates with bounded-variation amplitudes, applied after cutting away the endpoints and then passing to the limit. For $u = 0$ the first-derivative

estimate suffices. This proves (74). The matching lower bound on that norm follows from the $2n$-shell coefficient, of size comparable to $n^{\alpha/2}$ since $c_0 = 1$, and the shell $L^1(dx)$ bound above. Thus $Y_n = \|G_n\|_\infty \asymp n^{-1/2}$.

The oscillator orthogonality at a gives

$$\|\mathcal{T}_a G_n\|_2 \asymp \lambda_a^n n^{\alpha/2}, \qquad \|(\mathcal{N}+1)\mathcal{T}_a G_n\|_2 \le C n^{1+\alpha/2}\lambda_a^n.$$

For this upper bound tilt the signal on each shell k by $\exp(c/(k+1))$ and use the coarse Gram operator estimate at levels near a. Summing the resulting geometric tail gives the claim.

On the fixed annulus $1 \le u \le 2$, the fixed-argument Laguerre asymptotic [12, 13] gives

$$e^{-u/2}L_{n+j}^{(\alpha)}(u) = \frac{n^{\alpha/2-1/4}}{\sqrt{\pi}\,u^{\alpha/2+1/4}}\left(\cos(2\sqrt{nu} - \alpha\pi/2 - \pi/4) + o(1)\right)$$

uniformly for $j \le C\log n$. After applying $\tau_a(b)^{\mathcal{H}/2}$, the leading coefficient sum is $\sum_j c_j \tau_a(b)^j = (1-\tau_a(b))^l(1+\tau_a(b)) > 0$. The tail $j > C\log n$ is negligible by the same fixed-annulus bounds and the exponential factor, taking the fixed C large. The square of the oscillatory cosine integrates to a positive limit against any smooth positive weight on a smaller fixed annulus, by integration by parts. Using (69) on that annulus yields

$$\|\mathcal{T}_b G_n\|_2 \ge c\tau_a(b)^n n^{\alpha/2-1/4}.$$

Its component at Gaussian degrees $k \le \epsilon n$ is negligible for small fixed $\epsilon > 0$. Indeed its norm is at most $(b/b')^{\epsilon n/2}\|\mathcal{T}_{b'}G_n\|_2$ for any $a < b' < b$. The shell estimates bound the latter by $\tau_a(b')^n$ times a polynomial, strictly exponentially smaller after choosing ϵ small. Thus $\|\mathcal{N}\mathcal{T}_b G_n\| \ge cn\|\mathcal{T}_b G_n\|$ for the purpose of the same lower bound. Both logarithms in Theorem 11.1 are of order n on this sequence. Substituting these powers of n gives the two claimed sharp logarithmic exponents.

11.2.2 The critical constant at small reference levels

For completeness, let $m = 2$, let $b = b(a)$ solve $\tau_a(b)^2 = \lambda_a$, and let $d = 1, 2$. The best constant in $\|\mathcal{T}_b g\|_2^2 \le C_{a,d} YX$ has order $a^{-d/8}$ as $a \downarrow 0$. This settles the boundedness question in [5, Section 9.3, item 2] in dimension one.

Indeed $\tau = \tau_a(b) = \sqrt{\lambda_a} \asymp \sqrt{a}$, and the Gaussian coefficient c in (69) satisfies $c \asymp 1 - \nu_a(b)/\tau \asymp \tau$. The constants in the Gram factorizations at a, b are bounded above and below as $a \downarrow 0$. For the one-coordinate localization, write $\xi = \nu_a(b)/\tau$. The generating function after normalization has coefficients from $((1-z)(1-\xi z))^{-1/2}$. At shell k, their sum is at most one, by comparison with $\xi = 1$. Split the convolution at the index of ξ^j being $k/2$: the first part is at most $C((k+1)(1-\xi))^{-1/2}$ by the generating sum of that factor, and the second at most $\xi^{k/2}$. Consequently the shell bound (70) can be replaced uniformly by $C\min\{1, (c(k+1))^{-1/4}\}$.

Write $r_k = (Y/X)(k+1)^{d/4}\lambda_a^{k/2}$. At equality $\tau^2 = \lambda_a$, the contribution of the kth shell to the norm is at most

$$C\sqrt{YX}\,(k+1)^{d/8}\min\{1, (c(k+1))^{-1/4}\}\min\{r_k^{1/2}, r_k^{-1/2}\}.$$

For small a the successive ratios of r_k are uniformly bounded above by a constant less than one. Hence the sum of the last factors is uniformly bounded. For $d \le 2$ the supremum of the middle two factors is at most $Cc^{-d/8}$. This proves $C_{a,d} \le Cc^{-d/4}$.

For the lower bound use G_n above with $n \asymp 1/c$. The factors B_a in the norm comparisons stay in a compact set. After $\tau^{\mathcal{H}/2}$, the first radial mode dominates in $L^2(dx)$: the sum of the other modes

has norm at most $C\tau$ times the norm of the first, by the geometric factor τ^j. The Gaussian multiplier of (69) retains a fixed fraction of that first mode's squared norm. To see this, the expectation of $|x|^2$ under its normalized squared Lebesgue density is $O(n)$; Jensen's inequality bounds the expectation of $e^{-c|x|^2}$ below by e^{-Ccn}. Thus $\|\mathcal{T}_b G_n\|_2^2 \gtrsim \tau^{2n} n^\alpha$ uniformly here. Meanwhile $\|G_n\|_\infty \lesssim n^{-1/2}$ and $\|\mathcal{T}_a G_n\| \lesssim \lambda_a^n n^{\alpha/2}$ as above. The ratio is at least a constant times $n^{(\alpha+1)/2} = n^{d/4}$. This gives $C_{a,d} \gtrsim a^{-d/8}$.

11.3 A collective estimate for conditional products

Given m independent standard Gaussian vectors Z_i and independent Gaussian noise, let $\mathcal{B}_{q,w}$ be conditional expectation of $\prod_i f_i(Z_i)$ given the standard Gaussian output at signal q and weights w. The notation $\mathcal{B}_{q,w}(g)$ means that all factors are g. We only need the following estimate on compact positive levels, near the equal-weight configurations $w = u$, $a = q/m$, $b_i = q$.

Lemma 11.3. *Take q, w, a, b_i in a sufficiently small neighborhood of such a configuration, with q in a fixed compact subinterval of $(0,1)$. Suppose*

$$\sum_i q w_i / b_i = 1, \qquad \prod_i \tau_a(b_i) \le \lambda_a.$$

For bounded g, with $Y = \|g\|_\infty$ and $X_1 = \|(\mathcal{N}+1)\mathcal{T}_a g\|_2$, one has

$$\|\mathcal{N}\mathcal{B}_{q,w}(g)\|_2 + \|\mathcal{N}^{-1}\partial_v^2 \mathcal{B}_{q,w}(g)\|_2 \le C Y^{m-1} X_1$$

for every unit tangent v. The inverse number multiplier is zero on constants. The estimate is uniform on the indicated compact neighborhoods.

Proof. We first prove a shell bound for the conditional operator before summing over the inputs. In the Gram pairing for this operator, use two independently conditioned input replicas given the same output. Their covariance has identity diagonal blocks and cross block $qcc^\top \otimes I_d$, where $c_i = \sqrt{w_i}$. In oscillator generating variables z, z' the Gaussian integration gives the exponent

$$\tfrac{1}{4}(z+z')^\top P_{\mathrm{orig}}(z+z') - \tfrac{1}{4}(z-z')^\top Q_{\mathrm{orig}}(z-z'),$$

where matrices on copy indices act as the identity on spatial coordinates and

$$P_{\mathrm{orig}} = r_0 I + (\nu_a(q) - r_0) cc^\top, \qquad Q_{\mathrm{orig}} = -r_0 I + (\tau_a(q) + r_0) cc^\top, \qquad r_0 = \lambda_a^2.$$

For example the sum and difference replica covariances have eigenvalues $1+q, 1-q$ along c, and 1 transversely. Integrating the oscillator ground factors and their generating functions sends these eigenvalues to the indicated parameters, exactly as in the one-dimensional Gram calculation. The determinant prefactor is fixed and positive.

Write $\tau_i = \tau_a(b_i)$ and $D = \mathrm{diag}(\tau_i)$. Scale each input generating variable by $\tau_i^{-1/2}$, so the remaining matrices are $P = D^{-1/2} P_{\mathrm{orig}} D^{-1/2}$ and $Q = D^{-1/2} Q_{\mathrm{orig}} D^{-1/2}$. For shell degrees n_i, this removes the factor $\prod_i \tau_i^{n_i}$ from the squared norm. The relation

$$\tau_a(b) + r_0 = \frac{2 d_a b}{(1+d_a)(1+d_a-b)}$$

and the normalization of the b_i give $(\tau_a(q) + r_0) \sum_i w_i/(\tau_i + r_0) = 1$. Thus $I - Q$ is positive semidefinite with a one-dimensional null space on the copy indices. Let e be its unit null vector,

chosen continuously near equal weights. Also $0 < P < I$, and Q has other eigenvalues in $(-1, 1)$, by continuity from the equal configuration.

The normalized Gram has a useful physical Gaussian form. Let (U, V) be a pair of vectors in $(\mathbb{R}^d)^m$ with independent sum and difference coordinates $(U + V)/\sqrt{2}, (U - V)/\sqrt{2}$ of covariances $(I + P)/(2B_a), (I - Q)/(2B_a)$, respectively, with the spatial identities understood. Indeed its Hermite generating function after dividing out the oscillator ground factors is exactly the normalized Gram above. The collective coordinate $s = e^\top U = e^\top V$ agrees almost surely in the two replicas. Their cross covariance is $(P + Q)/(4B_a)$, of rank one on copy indices, because $P + Q = (\nu_a(q) + \tau_a(q))D^{-1/2}cc^\top D^{-1/2}$. Thus the two transverse vectors are independent given s: Gaussian regression subtracts exactly that rank-one cross covariance, since s is shared and has positive variance.

Let $\mathcal{U}$ denote rotation in copy space to these collective and transverse coordinates. For a finite oscillator tensor f, the normalized Gram quadratic form is therefore a positive constant times

$$\int p(s)e^{B_a|s|^2} \left| \int (\mathcal{U}f)(s, z)\, \phi_s(z)\, dz \right|^2 ds, \qquad \phi_s(z) = e^{B_a|z|^2/2} p_s(z),$$

where p is the Gaussian density of s, and p_s is the conditional transverse Gaussian density. At the equal configuration the collective variance is $(1 + \nu_a(q)/\tau_a(q))/(4B_a)$. The transverse covariance is $(1 + r_0/\tau_a(q))I/(2B_a)$ and its conditional mean is zero. Consequently the outer weight is bounded by $Ce^{-c|s|^2}$ with $c > 0$, and ϕ_s is a Gaussian with strictly positive quadratic decay. Both margins persist nearby, and the transverse mean is a linear function of s with arbitrarily small operator norm in a small neighborhood.

We can add a fixed transverse oscillator attenuation to this estimate. More precisely, for some $\sigma_0 < 1$ sufficiently close to one,

$$\|\sigma_0^{-\mathcal{H}_\perp/2}\phi_s\|^2_{L^2(dz)} \le C \exp(c'|s|^2)$$

with c' smaller than the decay in the outer weight. To verify this, expand the Gaussian ϕ_s by integrating it against the oscillator generating function. In the complex Gaussian polynomial representation, its series is a constant depending on s times $\exp(\zeta^\top A\zeta/2 + \ell_s^\top \zeta)$. The real symmetric matrix A has norm strictly below one, since its eigenvalues for a Gaussian with precision $C_0 > 0$ are those of $(B_aI - C_0)(B_aI + C_0)^{-1}$. The linear coefficient is $O(\epsilon|s|)$ and the prefactor is bounded by $Ce^{C\epsilon^2|s|^2}$, where ϵ tends to zero with the neighborhood. Applying $\sigma_0^{-\mathcal{H}_\perp/2}$ replaces ζ by $\zeta/\sqrt{\sigma_0}$. The squared complex Gaussian integral is finite with a uniform quadratic gap when $\|A\| < \sigma_0$. Completing the square gives the asserted bound. Cauchy–Schwarz in the transverse pairing now bounds the normalized Gram by a constant times

$$\|e^{-c_1|s|^2}\sigma_0^{\mathcal{H}_\perp/2}\mathcal{U}f\|^2_{L^2(ds\, dz)}$$

for a fixed $c_1 > 0$.

Take $f = \bigotimes_i f_{n_i}$, where f_{n_i} is in input oscillator shell n_i, and $N = \sum_i n_i$. Rotation preserves total oscillator degree. For transverse degree j, the collective degree is $N - j$, so (70) supplies $(N - j + 1)^{-1/4}$ in norm. Different transverse degrees are orthogonal in dz. Because $(N + 1)/(N - j + 1) \le j + 1$, we can absorb the difference between $N - j$ and N by replacing σ_0 by one fixed $\sigma \in (\sigma_0, 1)$. Restoring the input attenuation, we obtain

$$\|\mathcal{B}_{q,w}(f_{n_1}, \dots, f_{n_m})\|_2 \le C(N + 1)^{-1/4} \prod_i \tau_i^{n_i/2} \|\sigma^{\mathcal{H}_\perp/2}\mathcal{U}f\|_{L^2(dx)}. \tag{75}$$

The same bound with an additional factor $C(N+1)$ holds for one output number power. Enlarge q by $\exp(c_0/(N+1))$, enlarging every b_i by the same factor to preserve normalization. The collective null direction e and the transverse attenuator may change under this enlargement, but the bounds used below on the final tensor norm are uniform in that direction, and the attenuation can be chosen uniformly. The enlargement changes the product of the τ_i powers by a bounded factor. For $(\mathcal{N}+1)^{-1}$ the factor is instead $C/(N+1)$. Use its integrated OU formula as in (72), lowering q and the b_i by a common factor v. Near one the product of powers decreases by $e^{-cN(1-v)}$, and away from one use OU contraction. In these two weighted assertions, the tensor norms on the right can be replaced by their uniform estimate below before performing the enlargement or integration.

The uniform tensor norm. In the complex Gaussian polynomial representation, an oscillator shell vector corresponds to a homogeneous polynomial F_{n_i} of degree n_i and norm $\|f_{n_i}\|_{L^2(dx)}$. With independent standard circular complex vectors in the rotated coordinates, the last squared norm in (75) is $\mathbb{E}\prod_i |F_{n_i}(Z_i')|^2$. The covariance on copy indices of Z' is $\sigma I + (1-\sigma)ee^\top$. It is invariant under a common unitary spatial rotation. Write $C_i = \sup_{|\zeta|=1} \sqrt{n_i!}|F_{n_i}(\zeta)|$. Insert a common Haar unitary rotation, keeping the spherical average on one factor i_0 and taking the supremum on the others. The exact complex spherical average is

$$\int |\sqrt{n_i!}F_{n_i}(\zeta)|^2 \, d\zeta = \|f_{n_i}\|^2_{L^2(dx)} \Big/ \binom{n_i+d-1}{d-1}.$$

This follows from polar integration for the complex Gaussian. The remaining radial moment is $\mathbb{E}\prod_i(|Z_i'|^{2n_i}/n_i!)$. Pointwise its integrand is $V^N/N!$ times a multinomial probability, where $V = \sum_i |Z_i'|^2$. The probability is at most $C\sqrt{N+1}/\prod_i \sqrt{n_i+1}$ for all probability vectors, by maximizing at the empirical proportions and using Stirling's bounds, including zero indices. Also $\mathbb{E}V^N/N! \le C(N+1)^{d-1}$, because V has law $\mathrm{Gamma}(d,1) + \sigma\mathrm{Gamma}(d(m-1),1)$ with independent summands. Expanding its moments gives this bound directly, since the second summand has a geometrically summable weight.

Combining with (75), the factor $(N+1)^{-1/4}$ cancels the numerator from the multinomial mode. With $k = n_{i_0}$, the resulting bound for the unweighted output is

$$C\prod_i \tau_i^{n_i/2} \left(\frac{N+1}{k+1}\right)^{(d-1)/2} \prod_i (n_i+1)^{-1/4} \|f_{n_{i_0}}\|_{L^2(dx)} \prod_{i\neq i_0} C_i. \tag{76}$$

This bound is uniform in e. The same estimate has an extra factor $N+1$ or $(N+1)^{-1}$ for the two output weights just considered.

Coherent estimates on an input shell. In the oscillator space let A_j^* denote the creation operators. The normalized vector

$$\frac{(\zeta\cdot A^*)^n}{\sqrt{n!}} H_0, \qquad |\zeta| = 1,$$

represents evaluation $\sqrt{n!}F_n(\zeta)$, up to complex conjugation in the pairing. Its $L^1(dx)$ norm is at most $C(n+1)^{1/4}$, uniformly over complex unit vectors ζ. Indeed the generating vector $\exp(z\,\zeta\cdot A^*)H_0$ has physical form

$$H_0(x)\exp\left(\sqrt{2B_a}\, z\zeta\cdot x - z^2\zeta\cdot\zeta/2\right).$$

Completing the square in its absolute integral gives $C\exp(|z|^2/2)$. Cauchy's formula at $|z| = \sqrt{n}$, with the normalized coefficient $\sqrt{n!}$, gives the bound by Stirling. For $n = 0$ use the ground state directly. Thus for $f_n = P_n g$,

$$C_i \le C\min\{Y(n+1)^{1/4}, X_1(n+1)^{-1}\lambda_a^{-n/2}\}. \tag{77}$$

The second bound follows from (72) and the unit norm of the evaluation vector.

We explain how the weight derivatives fit this estimate. Differentiating the conditional Hermite coefficient in two weight directions gives a sum of fixed products of Gaussian input number multipliers $\mathcal{N}_i$, of total degree at most two in these multipliers. The coefficients remain bounded on positive weights away from zero. Each $\mathcal{N}_i$ is a fixed quadratic sum of Lebesgue oscillator creation and annihilation operators. A monomial of length l_0 maps an input shell n to a shell of degree $n+O(l_0)$ and has $L^2(dx)$ cost at most $C(n+1)^{l_0/2}$. It has the same cost in the coherent bounds (77). For the sup-norm input bound, move its adjoint onto the coherent vector. For a fixed monomial the vector then lies exactly in the original shell, so pairing with $P_n g$ is pairing with g. Applied to the generating Gaussian, the adjoint introduces a polynomial in z, x of degree at most l_0. Its absolute Lebesgue integral is bounded by $C(1+|z|)^{l_0}\exp(|z|^2/2)$ by the same shifted Gaussian calculation. Cauchy's formula gives the claimed factor. The energy-side bound follows from the $L^2(dx)$ monomial norm. All fixed degree shifts change the shell factors and attenuations by bounded constants. Thus the total cost of two weight derivatives, followed by $\mathcal{N}^{-1}$ on positive degrees, is at most one factor $C(N+1)$ in (76), using the inverse output weight already proved. One output number has exactly that same cost.

Summation. In (76) always choose i_0 with minimum input shell degree $k = \min_i n_i$. Set $l_1 = N - mk$. The degree shifts just discussed may be absorbed in constants, so the minimum can be chosen using the original degrees. Use the small energy bound for $\|P_k g\|$ in (72), and the sup bound in (77) for the other factors. Since $\prod_i \tau_i \le \lambda_a$ and $\max_i \tau_i$ is bounded away from one, the resulting summand for either derivative estimate is at most

$$CY^{m-1}X_1(k+1)^{-1/4}(l_1+1)^C\rho^{l_1}, \qquad \rho < 1. \tag{78}$$

Indeed the k-part of the exponential attenuation cancels $\lambda_a^{-k/2}$, and $(N+1)/(k+1)$ is at most $m + l_1$. The number of tuples of excess l_1 is polynomial, so it is absorbed in the exponential decay.

To sum the minimum degrees define K_1, K_d by

$$(K_1+1)^{5/4}\lambda_a^{K_1/2} \asymp X_1/Y, \qquad (K_d+1)^{1+d/4}\lambda_a^{K_d/2} \asymp X_1/Y.$$

Use the decreasing branch of each expression and bounded K_1, K_d when X_1/Y is bounded below. They agree up to $O(\log(K_1+2))$, and are comparable when large. On the interval between them the factor $(k+1)^{-1/4}$ in (78) pays for the logarithmic number of degrees. Above K_d use the large shell bound for $P_k g$ instead of the small bound. Their ratio is bounded by a polynomial times $\lambda_a^{(k-K_d)/2}$, so this gives a summable geometric gain. Below K_1, if any other degree satisfies $n_i \le (K_1+k)/2$, use its small coherent bound instead of its large one. This gives a geometric gain in $K_1 - k$, since the ratio of the bounds is at most

$$C\left(\frac{K_1+1}{n_i+1}\right)^{5/4}\lambda_a^{(K_1-n_i)/2}.$$

If every other degree is above that midpoint then $l_1 \ge (K_1-k)/2$, giving the same gain from ρ^{l_1}. Thus all sums are bounded by $CY^{m-1}X_1$.

All estimates were stated on finite shells. For bounded g first cut off its support and use finite oscillator sums. Oscillator convergence in $L^2(dx)$ implies $L^2(\gamma_d)$ convergence, and the conditional operator and its fixed weight derivatives are continuous on the Gaussian tensor space after positive smoothing. Equivalently their coefficients are bounded by tensor contraction at a slightly enlarged total signal, which pays the finite number powers. The absolutely convergent shell estimates pass to these limits. Cutting off at increasing radius then converges in $L^2(\gamma_d)$, and X_1 also converges since $(\mathcal{N}+1)\mathcal{T}_a$ is bounded there. This proves the lemma. □

11.4 Entropy curvature in dimension d

The output density relative to γ_d is $1 + R = 1 + L + M$, where $L = \sum_i \mathcal{T}_{qw_i} g$ and M is the sum of the conditional products on subsets of at least two factors. For multi-indices with $\alpha = \sum_i \alpha_i$ on a subset, the output coefficient summands are

$$\left(\frac{\alpha!}{\prod_i \alpha_i!}\right)^{1/2} \prod_i (qw_i)^{|\alpha_i|/2} \widehat{g}(\alpha_i).$$

Factorials here are coordinatewise. Every active total input degree is even and at least four, by central symmetry, mass and covariance. All linear weight multipliers depend only on total degree.

For a real tangent direction set $E = \langle L'', L\rangle$. At degree $2k$, L and L'' have multipliers $q^k \sum_i w_i^k$ and $q^k k(k-1) \sum_i v_i^2 w_i^{k-2}$. In particular $\|\mathcal{N}^{-1} L''\|_2 \le C E^{1/2}$ on compact positive weights. If the tangent has a fixed positive fraction of its norm on a largest weight, then $E \ge c\|\mathcal{N} L\|_2^2$. This holds on every ray $w = u + sv,\ s \ge 0$, because a largest weight coordinate has $v_i = \max_j v_j$, bounded below by a positive multiple of $\|v\|_2$ on the tangent space. For two copies it also holds at every weight including endpoints: the ratio of the two multipliers in the coordinate $t = w_1$ lies between $k(k-1)$ and $8k(k-1)$, because $\max(t, 1-t) \ge 1/2$. These scalar estimates count Gaussian degrees and are independent of spatial dimension.

Put $\delta = \|g\|_\infty$. Gaussian conditioning bounds the spatial derivatives of L, M, R through order two and their sup norms by $C\delta$ on compact signal sets below one. Indeed the conditional covariance in input space is bounded below by $1 - q$; derivatives of its mean kernel introduce fixed Gaussian polynomials. Where the weight derivatives exist in L^2, the entropy formula is

$$F'' = -\langle L'' + M'', L + M + \psi(R)\rangle - \int \frac{(R')^2}{1+R}\, d\gamma_d, \qquad \psi(z) = \log(1+z) - z.$$

This follows by differentiating relative entropy, since the output has mean zero and fixed covariance I_d, and $1 + R$ stays bounded away from zero. The formula also holds by continuous one-sided extension at a weight endpoint if the corresponding L^2 derivatives extend there. The number chain rule satisfies

$$\|\mathcal{N}\psi(R)\|_2 \le C\delta \|\mathcal{N} R\|_2. \tag{79}$$

Indeed Gaussian integration by parts against $R|\nabla R|^2 \nabla R$ bounds $\||\nabla R|^2\|_2$ by $3\|R\|_\infty \|\mathcal{N} R\|_2$, since $\|\nabla^2 R\|_{L^2(\mathrm{HS})} \le \|\mathcal{N} R\|_2$ by Hermite expansion. This and $\mathcal{N}\psi(R) = \psi'(R)\mathcal{N} R - \psi''(R)|\nabla R|^2$ give (79).

We will also use the strict oscillator L^P estimates from Sections 9.1–9.5 in dimension d. Tensoring the one-dimensional mode bounds for $b > a$ and $P > 2$ gives

$$\|\mathcal{T}_b H_\alpha\|_P \le C_d z^{|\alpha|/2} \quad \text{for} \quad z > \max\{\tau_a(b), \sigma_a(b, P)\},$$

where

$$\sigma_a(b, P) = \frac{B_a[1 + (P-1)b] - 1}{B_a[1 + (P-1)b] + 1}.$$

For tensor oscillator coefficients the Lebesgue L^1 bound is polynomial in $|\alpha| + 1$, and the other coefficient bound is $C\|\mathcal{T}_a g\|_2 \lambda_a^{-|\alpha|/2}$. The count of multi-indices of fixed total degree is polynomial. Interpolating the coefficient bounds and using the strict exponential margin yields

$$\|\mathcal{T}_b g\|_P \le C\delta^{1-p} D_a^p, \qquad D_a = \|\mathcal{T}_a g\|_2, \qquad 0 < p < \frac{\log \max(\tau_a(b), \sigma_a(b, P))}{\log \lambda_a}. \tag{80}$$

Theorem 11.1 gives the $P = 2$ estimate for every $p < \log \tau_a(b)/\log \lambda_a$. At $b = a$ and $P = 2$ both parameters equal λ_a. Writing $z_a = \sqrt{(1-a)/(1+a)}$, one has $\partial_P \log \sigma_a(a, 2) = z_a$ and $\log(1/\lambda_a) = 2\,\mathrm{artanh}(z_a) > 2z_a$. Thus for $P > 2$ close to two and $b > a$ close to a, one may take $pP/2 > 1$ in (80), uniformly on compact reference levels.

Here is a consequence for the logarithm in an off-centre Hessian, used when $\|\mathcal{N}L\|_2^2$ need not be controlled by E. If all linear levels lie sufficiently close to a_{ref} and admit a slight enlargement with the margin just described, then $\|\mathcal{N}\psi(L)\| \le C\delta D_{a_{\mathrm{ref}}}$. Factor L through that enlarged level, leaving a strictly positive residual OU smoothing for its first two spatial derivatives. Gaussian kernel differentiation and Hölder control those derivatives from L^P to L^P. Interpolate each bound in (80) with its $C\delta$ sup bound to an exponent $U > 4$ sufficiently close to four that $2Pp/U > 1$. Products of two of the resulting bounds give $C\delta D_{a_{\mathrm{ref}}}$, including in terms with a factor $|x|$ by its finite Gaussian L^A norm where $1/A + 2/U = 1/2$. The number chain rule gives the claim. For sufficiently small q, with all levels of order q, the same proof follows instead from Gaussian hypercontractivity [11] and spectral Hölder, for example using $P = 3$, $p = 3/4$, since an enlarged level of order q to power $1/p$ is then below a_{ref}.

For the difference $W = \psi(R) - \psi(L)$, the chain rule gives

$$\|\mathcal{N}W\|_2 \le C\delta \|(\mathcal{N}+1)M\|_2. \tag{81}$$

Indeed subtract the two number-chain rules. The terms containing $\mathcal{N}M$ or ∇M have coefficients $O(\delta)$. The other terms contain M times $\mathcal{N}L = O(\delta(1+|x|))$ or $|\nabla L|^2 = O(\delta^2)$. The Gaussian Hermite ladder bounds $\|(1+|x|)M\|_2 + \|\nabla M\|_2$ by a constant times $\|(\mathcal{N}+1)M\|_2$.

11.4.1 Two copies, including the endpoint

It is enough to prove $q = 4/7$. At smaller q apply a common input OU smoothing at signal $q/(4/7)$ before the endpoint experiment. This preserves admissibility and contracts the relative sup norm, and the resulting entropy equals the desired entropy at signal q. The input OU operator is injective on $L^2(\gamma_d)$, so non-Gaussian inputs remain non-Gaussian.

Near $t = 1/2$, take $a = q\max(t, 1-t)$ and $b_1 = b_2 = q$. Then $\tau_a(q)^2 \le \lambda_a$ since it holds at $a = q/2$, and $\tau_a(q)$ decreases and λ_a increases with a. Lemma 11.3 gives $\|\mathcal{N}M\| + \|\mathcal{N}^{-1}M''\| \le C\delta\|\mathcal{N}L\|$. Together with (79), every error in the entropy pairing is $O(\delta E)$. Here we use the linear coercivity above and pair second derivatives through $\mathcal{N}^{-1}$.

Away from a fixed small interval about $1/2$ there is a strict gap $\tau_a(q)^2 < \lambda_a$ uniformly. Use a small common complex disk at each real weight t. For the mixed subset the total complex signal is at most $q(|z| + |1-z|)$, and can be uniformly enlarged to a level b with $\tau_a(b)^2 < \lambda_a$. Absolute weights can be used in Gaussian conditional contraction after applying unit Hermite phases to each input; the input OU norms are unchanged. All weight exponents are integers even at 0 and 1. The strict estimate at b and the extra attenuation give an analytic Gaussian L^2 series, also with any fixed number powers on a smaller disk. By Cauchy, $\|\mathcal{N}M\| + \|M''\| \le C\delta\|\mathcal{T}_a g\|_2$. This proves the same entropy bound and continuity at the weight endpoints. The conditional real Gaussian formula supplies the bounded positive density ratio throughout the closed weight interval. Taking δ small gives the strict curvature. Luo's one-dimensional counterexamples for $q > 4/7$ extend by adding independent Gaussian coordinates.

11.4.2 More copies

The strict threshold algebra of Sections 9.1–9.5 is dimension-free. At $w = u$, $a = q/m$, a size-j subset has signal ja, and $\tau_a(ja)^j < \lambda_a$ for every $2 \le j \le m$ when $q < q_m$. The margins for proper

subsets remain strict at $q = q_m$. Fix $Q < q_m$ and first restrict q away from zero. For w near u take $a_{\mathrm{ref}} = q(1/m - \xi)$ with $\xi > 0$ small. The linear energy controls $\|(\mathcal{N}+1)\mathcal{T}_{a_{\mathrm{ref}}} g\|_2^2$, by keeping a largest weight in L and the smallest weight in its second derivative. On fixed complex disks about these weights all subset signals, even after removing a small fixed output attenuation $\mathcal{T}_{p_0}$, can be enlarged to levels b_j with $\tau_{a_{\mathrm{ref}}}(b_j)^j < \lambda_{a_{\mathrm{ref}}}$. We choose $p_0 < 1$ close to one and shrink the ball so that $p_0 \max_i w_i \le a_{\mathrm{ref}}/q$. The strict L^2 comparison, tensor contraction and Cauchy then bound M, M'' after removing $\mathcal{T}_{p_0}$ by $C\delta D_{a_{\mathrm{ref}}}$. This controls $\langle M'', L\rangle$ by transferring the attenuation, and controls $\langle L'', M\rangle$ using the inverse number pairing. The remaining logarithmic pairings use $\|\mathcal{N}\psi(L)\| \le C\delta D_{a_{\mathrm{ref}}}$ and (81). They are all $O(\delta E)$. For q small enough one uses the same argument with spectral Hölder in the subset estimates, since every enlarged size-j signal is $O(q)$ and its jth power is below a_{ref}. This proves the common ball below q_m.

At q_m, along a ray near u take $a = q_m \max_i w_i$ and $b_i = q_m$. Then the full subset satisfies Lemma 11.3, and proper mixed subsets still have strict oscillator margin. The linear energy controls $\|\mathcal{N}L\|_2^2$, and the product estimate gives $\|\mathcal{N}M\| + \|\mathcal{N}^{-1}M''\| \le C\delta\|\mathcal{N}L\|$. Use (79) and number duality as in the two-copy proof. This gives radial curvature, including the exact barycentric Hessian.

For the full endpoint ball when $m = 3, 4$, use the weight geometry of Sections 10.1–10.3. Order the weights increasingly and set

$$a = \min\{q_m/m, q_m\sqrt{w_{(2)}w_{(m)}}\}.$$

The linear energy controls $X_1^2 = \|(\mathcal{N}+1)\mathcal{T}_a g\|_2^2$: at least $1/m$ of any unit tangent's squared norm lies on indices other than the smallest, and a largest weight occurs in the linear factor L. For three copies $a = q_m/m$ near u because $w_{(2)}w_{(3)} \ge 1/9$, so take $b_i = q_m$. For four copies take $b_i = q_m\sqrt{w_i}\sum_j \sqrt{w_j}$. The normalization $\sum_i q_m w_i/b_i = 1$ is immediate. The scalar comparison of Sections 10.1–10.3, obtained by concavity of $\log\tau_a$ and Taylor expansion, gives $\prod_i \tau_a(b_i) \le \lambda_a$ near u. Indeed $a \ge (q_m/4)(1 - 4\|w-u\|^2 + O(\|w-u\|^4))$ and the mean of the b_i is $q_m(1 - \|w-u\|^2 + O(\|w-u\|^3))$. The gain from lowering their mean exceeds the loss from lowering a, with derivative comparison of sign $q_m^2(9+6q_m)/16 > 0$. Thus Lemma 11.3 gives the needed number bounds for the full subset in both cases: $\|\mathcal{N}M\| + \|\mathcal{N}^{-1}M''\| \le C\delta X_1$, with the proper subsets following strictly.

For the two linear–mixed pairings at this endpoint, a strict transfer is still available. Apply $\mathcal{T}_{p_0}$ to the mixed factor and its inverse to the linear factor, with $p_0 < 1$ close to one. At the centre and reference $a_0 = q_m/m$, the available exponents of $\|\mathcal{T}_{a_0} g\|$ for the full and linear factors are

$$m\frac{\log\tau_{a_0}(p_0 q_m)}{\log\lambda_{a_0}}, \qquad \frac{\log\tau_{a_0}(a_0/p_0)}{\log\lambda_{a_0}}.$$

Their sum is two at $p_0 = 1$. It exceeds two for $p_0 < 1$ sufficiently close to one because

$$mq_m\partial_b \log\tau_{a_0}(q_m) > a_0\partial_b\log\tau_{a_0}(a_0).$$

Proper subsets already have a strict margin. The strict sum persists at a reference a_{ref} slightly below a_0 and for w close to u, with $a_{\mathrm{ref}} \le a$. Part of the margin pays the number factors from weight derivatives. Conditional contraction and the strict scalar estimates thus give a product bound $C\delta D_{a_{\mathrm{ref}}}^2$ in each mixed pairing. The logarithmic error again uses the bound for $\mathcal{N}\psi(L)$ and (81). All errors are $O(\delta E)$, which proves the full ball at q_m for three and four copies. Input OU reduction gives the same ball below q_m.

For $m \ge 5$ the endpoint skew obstruction of Sections 10.1–10.3 lifts to every dimension by letting the perturbation depend only on the first coordinate. The same lift applies above q_m to the positive barycentric curvature examples of Sections 9.1–9.5. The remaining coordinates are

independent standard Gaussians before and after mixing, adding a constant to the entropy. The lift preserves covariance, arbitrary fixed Gaussian moment matching, smoothness, strong log-concavity, and all the relative supremum derivative convergences. This completes Theorem 11.2.

12 Dimension-uniform interpolation and entropy neighborhoods

12.1 The dimension-uniform interpolation power

Let γ_d be standard Gaussian measure in dimension d, and let $\mathcal{T}_b = T_{\sqrt{b}}$ have multiplier $b^{n/2}$ on Gaussian Hermite degree n. Write $\mathcal{N}$ for total Gaussian number, so its multiplier is n. Fix $0 < a < b < 1$ and set $\theta_0 = \log b / \log a$. For a bounded function g write $Y = \|g\|_\infty$, $X = \|\mathcal{T}_a g\|_2$, with norms for γ_d.

Theorem 12.1. *In every dimension,*

$$\|\mathcal{T}_b g\|_2^2 \le Y^{2-2\theta_0} X^{2\theta_0}.$$

For any $\theta > \theta_0$, the optimal constant for this inequality with θ in place of θ_0 is at least $\exp(cd)$ *in every sufficiently large dimension. The same obstruction holds after adding an arbitrary fixed power of* $1+\log(Y/X)$ *on the right. It holds for smooth bounded centrally even perturbations g with prescribed finite Gaussian moment matching.*

Here moment matching means $\int x^\alpha g \, d\gamma_d = 0$ for all multi-indices through the prescribed total degree. The stronger fixed dimension exponent in Sections 11.1–11.4.2, following the oscillator comparisons in [7, 5], involves both levels through a multiplier arising from Lebesgue oscillator modes. Theorem 12.1 determines the uniform power.

Proof. Let A_n be the squared $L^2(\gamma_d)$ norm of the Gaussian Hermite component of degree n. Since $b = a^{\theta_0}$, Hölder gives

$$\sum_n b^n A_n \le \left(\sum_n A_n \right)^{1-\theta_0} \left(\sum_n a^n A_n \right)^{\theta_0}.$$

Bound $\|g\|_2$ by Y.

For sharpness fix an odd integer n and a block of D independent standard Gaussian coordinates. Put

$$V_D(x) = D^{-1/2} \sum_{j=1}^{D} \mathrm{He}_n(x_j)/\sqrt{n!}, \qquad f_D(x) = \mathrm{sgn}(V_D(x)).$$

The zero set of V_D is null. The function is odd, has mean zero, and has sup norm one. If x, y are standard Gaussian blocks with coordinate cross-correlation c, then the summands in the two-dimensional vector $(V_D(x), V_D(y))$ have covariance matrix with unit variances and correlation c^n. This follows from the Hermite generating function. The ordinary bivariate central limit theorem yields convergence to that Gaussian pair, for each fixed n. The sign is continuous away from the two axes, which have zero limiting probability. Thus

$$\|\mathcal{T}_c f_D\|_2^2 \longrightarrow \frac{2}{\pi} \arcsin(c^n).$$

The sign expectation for a correlated Gaussian pair follows by expressing it as two projections of a rotationally invariant two-dimensional Gaussian.

For any fixed even number r of independent blocks put $g = f_D^{\otimes r}$. It is centrally even. If a monomial has total degree less than r, at least one block is absent from that monomial, so its integral against g is zero. By choosing r large enough we get any fixed prescribed moment matching. Both the sup norm and the squared noise stabilities tensor. For $\theta > \theta_0$, the limiting ratio at levels a, b for one block diverges along large odd n, since $\arcsin(c^n) \sim c^n$ and $b/a^\theta > 1$. Choose a fixed such n, then a finite D so that the ratio for g is strictly larger than one.

A smooth witness with the same strict inequality can be obtained on this fixed block. Approximate He_n by an odd compactly supported smooth truncation, then use $\tanh(KV_D)$ with the truncated summands and K large. These approximate the sign in Gaussian L^2, remain odd and bounded by one, and have bounded derivatives of all fixed orders. Gaussian contraction gives convergence of both noise stabilities. The product over the even number of blocks still matches the prescribed moments. Fix one resulting witness g_* of dimension d_* with ratio $R_* > 1$.

Take $g_*^{\otimes l}$ and add unused Gaussian coordinates if necessary. All norm ratios tensor, giving R_*^l. The logarithm $\log(Y/X)$ of the tensor power is l times the logarithm for the witness, so no fixed power of it absorbs the exponential increase. Taking $l = \lfloor d/d_* \rfloor$ proves the claim for every large d. □

The smooth witnesses may be multiplied by arbitrarily small positive numbers without changing any interpolation norm ratio. Since their first two derivatives are bounded, sufficiently small multiples also yield positive strongly log-concave density perturbations of γ_d. This observation concerns the scalar interpolation estimate.

12.2 A weighted conditional estimate

Theorem 12.2. *Fix $m \geq 2$ and $0 < a < b < 1$. A dimension-independent constant in*

$$\|\mathcal{N}\mathcal{T}_b g\|_2 \|\mathcal{T}_b g\|_2^{m-1} \leq C\|g\|_\infty^{m-1} \|(\mathcal{N}+1)\mathcal{T}_a g\|_2$$

exists exactly when $b^m \leq a$. If $b^m > a$, the same failure holds with any fixed power of $1 + \log(Y/X)$ on the right, even under the symmetry, smoothness and moment conditions of Theorem 12.1.

For positive weights w_i summing to one and independent standard Gaussian vectors Z_i, G, denote by $B_H(g)$ the conditional expectation of $\prod_{i\in H} g(Z_i)$ given $\sqrt{q}\sum_i \sqrt{w_i} Z_i + \sqrt{1-q}G$. Write $j = |H|$, $b = q\sum_{i\in H} w_i$. It equals the Gaussian conditional projection of an independent tensor of j factors $\mathcal{T}_b g$, with normalized weights on H. At fixed w this assertion follows by Hermite generating functions, and the projection contracts L^2 and identifies total input number with output number.

Lemma 12.3. *Suppose $0 < c_i \leq b_0 < 1$, $\sum_{i\in H} q w_i / c_i = 1$, and $\prod_{i\in H} c_i \leq a$. Then*

$$\|\mathcal{N} B_H(g)\|_2 \leq C_{j,b_0} Y^{j-1} \|(\mathcal{N}+1)\mathcal{T}_a g\|_2.$$

In a compact positive weight region the same bound holds for $\mathcal{N}^{-1} B_H''(g)$ along a unit tangent, with an additional weight-region constant. Both constants are independent of spatial dimension.

Proof. The coefficient of output Gaussian multi-index $\alpha = \sum_i \alpha_i$ before differentiation is a sum of terms

$$\left(\frac{\alpha!}{\prod_i \alpha_i!}\right)^{1/2} \prod_{i\in H} (q w_i)^{|\alpha_i|/2} \widehat{g}(\alpha_i).$$

A second weight derivative introduces a factor at most $C(\sum_i |\alpha_i|)^2$ in absolute value, for weights bounded below. After division by the positive output degree its cost before conditional contraction

is at most one tensor total number. The inverse is set to zero at degree zero. Freeze c_i at the given evaluation point and take the projection weights to be qw_i/c_i. The derivatives act as diagonal multipliers on the input tensor degrees before this frozen projection.

If n_i are the input degrees, let $k = \min_i n_i$ and $l = \sum_i n_i - jk$. Then

$$\left(\sum_i n_i\right)^2 \prod_i c_i^{n_i} \le C_{j,b_0}(k+1)^2 \left(\prod_i c_i\right)^k \le C_{j,b_0}(k+1)^2 a^k,$$

using the boundedness of $(l+1)^2 b_0^l$. Sum this inequality against products of the squared Gaussian degree norms. Taking the index of the minimum in turn bounds the sum by $CY^{2j-2}\|(\mathcal{N}+1)\mathcal{T}_a g\|_2^2$. Taking square roots proves both claims. □

Proof of Theorem 12.2. For the upper bound use the degree estimate just proved with $j = m$ and $c_i = b$, without applying a conditional projection. The squared norm with n_1^2 on the input tensor degrees is the squared left-hand side, and $n_1 \le \sum_i n_i$.

For necessity take the smooth mean-zero witness g_* from Theorem 12.1 at $\theta = 1/m > \theta_0$. Thus $\|\mathcal{T}_b g_*\|_2^m > \|g_*\|_\infty^{m-1}\|\mathcal{T}_a g_*\|_2$. For its l-fold tensor power the numerator dominates $\|\mathcal{T}_b g_*\|_2^{ml}$, since all active degrees are positive. The triangle inequality applied to total number gives

$$\|(\mathcal{N}+1)\mathcal{T}_a(g_*^{\otimes l})\|_2 \le (1 + lr_a)\|\mathcal{T}_a g_*\|_2^l, \qquad r_a = \frac{\|\mathcal{N}\mathcal{T}_a g_*\|_2}{\|\mathcal{T}_a g_*\|_2}.$$

Moment matching is preserved under tensor products and under adding unused coordinates. The numerator-to-right-side ratio grows exponentially, up to the harmless polynomial from the number factor and the optional logarithm. □

12.3 Dimension-uniform entropy neighborhoods

Call g admissible if it is bounded, centrally even, and $(1+g)\gamma_d$ has mass one and covariance I_d. For independent $X_i \sim (1+g)\gamma_d$ and independent standard Gaussian G, define

$$F_{q,g}(w) = h\left(\sqrt{q}\sum_i \sqrt{w_i}\, X_i + \sqrt{1-q}G\right).$$

Use the ordinary entropy for this density. Let $u = (1/m, \dots, 1/m)$.

Theorem 12.4. *There is an absolute Gaussian-relative sup norm radius on which two-copy entropy has strictly negative second weight derivative at every weight, including the continuous one-sided endpoints, for all $0 < q \le 1/2$ and in every spatial dimension, except at $g = 0$. For every fixed $m \ge 3$, put $q_0 = m^{-1/(m-1)}$. There are positive relative-density and weight radii depending only on m such that the radial curvature along every ray from u is strictly negative uniformly for $0 < q \le q_0$ and in every spatial dimension, except at $g = 0$. More generally, for each $1/2 < Q < 1/\sqrt{2}$ there is a radius depending only on Q giving the same conclusion for two copies whenever $0 < q \le Q$ and $4q\min(t, 1-t) \le 1$.*

The fixed-dimensional theorem of Sections 11.1–11.4.2 gives the larger exact two-copy interval $q \le 4/7$ and the larger multicopy thresholds from Sections 9.1–9.5 and 10.1–10.3. Theorem 12.4 supplies dimension-independent radii in its stated ranges. The scalar obstruction of Theorem 12.1 shows that above $1/2$ no scalar smoothing bound alone yields a dimension-independent entropy radius. This extends the local entropy-concavity theory of Section 2.

Proof. For the all-weight and radial claims it suffices to work at the largest signal. At lower q, replace the input perturbation by its OU smoothing at the ratio of the two signals. This preserves admissibility and contracts the sup norm. The output law is exactly the desired law at signal q, and a nonzero perturbation remains nonzero by Hermite injectivity.

Write $1 + R = 1 + L + M$ for the Gaussian-relative output, where $L = \sum_i \mathcal{T}_{qw_i} g$ and $M = \sum_{|H|\geq 2} B_H(g)$. The scalar linear multipliers in degree $2k$ are $q^k \sum_i w_i^k$ for L and $q^k k(k-1) \sum_i v_i^2 w_i^{k-2}$ for its second derivative. All active degrees are even and at least four. Put $E = \langle L'', L\rangle$ and $X_* = \|\mathcal{N}L\|_2$. For two copies the ratio of the multipliers in the coordinate $w = (t, 1-t)$ is between $k(k-1)$ and $8k(k-1)$, uniformly including endpoints. Along any radial direction near u, a largest weight has a direction component bounded below in absolute value by a constant times the norm of the tangent. Thus in the two respective regimes $E \geq cX_*^2$ and $\|\mathcal{N}^{-1}L''\| \leq CX_*$, with constants independent of dimension.

For two copies near $t = 1/2$, and for m copies near u, use $a = q \max_i w_i$ in Lemma 12.3. Here we take $c_i = b$ for each active index. For $m = 2$ at $q = 1/2$, $q^2 \leq a$ at every weight. For $m \geq 3$ at $q = q_0$, the full subset has $q_0^m = q_0/m \leq a$. At u every proper subset has strict $(q_0 j/m)^j < q_0/m$, $2 \leq j < m$. Indeed $j \log(q_0 j/m)$ is strictly convex in j, with equal endpoint values at 1 and m. The strict margins persist on a common ball. Since $\|(\mathcal{N}+1)\mathcal{T}_a g\|_2 \leq CX_*$ by the nonnegative linear multipliers,

$$\|\mathcal{N}M\|_2 + \|\mathcal{N}^{-1}M''\|_2 \leq C\|g\|_\infty X_*.$$

For two copies away from $1/2$, including weights $0, 1$, the inequality $q^2 < a$ is strict with uniform margin. On a fixed small complex disk about each such weight, enlarge the signal $q(|z| + |1-z|)$ to a common larger level b' with $(b')^2 < a$ at the real centre of the disk. The Hermite coefficients are polynomials in z and $1-z$, because the total input degrees are even. For absolute weights the coefficient norm is controlled by conditional contraction after unit phases are put into the input Hermite coefficients. Spectral Hölder gives $\|\mathcal{T}_{b'} g\|_2^2 \leq \|g\|_\infty \|\mathcal{T}_a g\|_2$. The remaining signal margin pays the number factor and gives a common analytic Gaussian L^2 bound. Cauchy's estimate gives the same bounds for M and M'' used above, now including the weight endpoints. All constants depend only on the fixed signal margin and disk radii, not on d.

For the extended two-copy range take $1/2 \leq q \leq Q$. When $\min(t, 1-t) \leq \eta$, for a sufficiently small fixed $\eta > 0$, we have the uniform strict margin $q^2 < q \max(t, 1-t) = a$. The preceding disk proof applies. On the remaining compact weight region take $c_1 = 2qt,\ c_2 = 2q(1-t)$. The criterion $4q \min(t, 1-t) \leq 1$ gives $c_1 c_2 \leq a$. The larger of these levels is below one on a further part of the range, as follows. Let $\Delta = (1 - 2Q^2)/4 > 0$. If $\max(t, 1-t) \geq q + \Delta$, use the disk proof again with the uniform strict margin. Otherwise the larger level is bounded by $2Q(Q + \Delta) < 1$. Lemma 12.3 and the same linear coercivity then give the required derivative estimates.

Finally, conditioning gives $\|R\|_\infty \leq (1 + \|g\|_\infty)^m - 1$. The entropy Hessian is

$$F'' = -\langle L'' + M'', L + M + \psi(R)\rangle - \int \frac{(R')^2}{1+R}\, d\gamma_d, \qquad \psi(z) = \log(1+z) - z.$$

Its continuous extension at $0, 1$ follows from the analytic Gaussian L^2 series and the bounded positive density ratio. In the spatial variable the Gaussian conditional kernel gives sufficient smoothness. The dimension-free Gaussian number chain rule is

$$\|\mathcal{N}\psi(R)\|_2 \leq C\|R\|_\infty \|\mathcal{N}R\|_2.$$

To see this, Gaussian integration by parts gives $\||\nabla R|^2\|_2 \leq 3\|R\|_\infty \|\mathcal{N}R\|_2$. Integrate the Gaussian divergence of $R|\nabla R|^2 \nabla R$ and use $\|\nabla^2 R\|_{L^2(\mathrm{HS})} \leq \|\mathcal{N}R\|_2$. The chain rule for $\mathcal{N}$ then gives the assertion. Pair second weight derivatives through $\mathcal{N}^{-1}$ in the entropy formula. All errors are

$O(\|g\|_\infty X_*^2)$, while the linear term is $-E \le -cX_*^2$. A sufficiently small radius, independent of dimension, proves both curvature claims. □

13 Dimension-uniform curvature at unequal weights

13.1 Statement

Let X_1, X_2 be independent with common density $(1+g)\gamma_d$, where γ_d is standard Gaussian measure, g is bounded and centrally even, and the density has mass one and covariance I_d. With an independent standard Gaussian G, put

$$F_{q,g}(t) = h\Big(\sqrt{qt}X_1 + \sqrt{q(1-t)}X_2 + \sqrt{1-q}G\Big), \qquad 0 \le t \le 1.$$

Write $r = q\max(t, 1-t)$ and $s = q\min(t, 1-t)$. Sections 2–5 give scalar local entropy restoration, and the dimension-uniform interpolation of Sections 12.1–12.3 treats all weights through signal $1/2$ and some outer weights above it. Luo's symmetric counterexamples for $q > 4/7$ can have greater entropy at an unequal weight than at the midpoint [7]. That comparison concerns global concavity and does not exclude negative curvature on outer weight intervals.

Theorem 13.1. *For each $Q < 1$, there is a positive number ε_Q which is independent of d. For every admissible nonzero g with $\|g\|_\infty < \varepsilon_Q$, every $0 < q \le Q$, and every weight satisfying*

$$\begin{cases} s \le 1/4, & r \le 1/2, \\ s \le r(1-r), & r \ge 1/2, \end{cases} \tag{1}$$

the curvature $F''_{q,g}(t)$ is strictly negative. At $t = 0, 1$ it is the continuous one-sided extension.

For $q \le 1/2$ this covers all weights. For $1/2 < q \le 3/4$, the condition is $\min(t, 1-t) \le 1/(4q)$. For $3/4 \le q < 1$, it is $\min(t, 1-t) \le \sqrt{1-q}/(1+\sqrt{1-q})$.

13.2 A conditional estimate with one unsmoothed factor

Let $\mathcal{T}_c = T_{\sqrt{c}}$ have Hermite multiplier $c^{n/2}$ in degree n, and let $\mathcal{N}$ have multiplier n. Write P_u for the conditional projection from the product Gaussian space to $\sqrt{u}Z_1 + \sqrt{1-u}Z_2$ for independent standard Gaussian vectors.

Lemma 13.2. *Fix $u_0, b_0 < 1$. Suppose $0 < u \le u_0$, $0 < c_1 \le 1$, $0 < c_2 \le b_0$, and $c_1 c_2 \le a < 1$. Let $\mathcal{D}$ be a diagonal Hermite multiplier whose absolute value is at most $n_1 + n_2$ on each block of input degrees n_1, n_2. Then*

$$\|P_u \mathcal{D}\,(\mathcal{T}_{c_1} g \otimes \mathcal{T}_{c_2} g)\|_2 \le C_{u_0,b_0}\|g\|_\infty \|(\mathcal{N}+1)\mathcal{T}_a g\|_2. \tag{2}$$

At $c_1 = 1$, the expression is defined by convergence after projection of the finite total-degree truncations.

Proof. Fix total degree N. The adjoint P_u^* maps a normalized Hermite polynomial of multi-index α to its expansion in the two input variables at their Gaussian linear combination. In the block of individual degrees $(n_1, N-n_1)$, its squared norm is

$$\binom{N}{n_1} u^{n_1}(1-u)^{N-n_1},$$

by $\sum_{|\beta|=n_1} \binom{\alpha}{\beta} = \binom{N}{n_1}$. The expansions of distinct α have disjoint multi-index supports even after restriction to these blocks. Consequently the squared operator norm of the projection on a union of such input blocks is exactly the corresponding binomial probability.

Choose τ with $u_0 < \tau < 1$. On blocks with $n_1 \le \tau N$, write $k = \min(n_1, n_2)$. Then $N \le n_2/(1-\tau)$, and

$$N^2 c_1^{n_1} c_2^{n_2} \le C(k+1)^2 a^k.$$

Use $(c_1 c_2)^k \le a^k$, $c_1 \le 1$, and $(1 + n_2 - k)^2 b_0^{n_2-k} \le C$. On the remaining blocks the squared projection norm at total degree N is at most e^{-cN}, with $c > 0$ depending on u_0, τ. This follows from exponential Markov applied to the binomial law with parameter at most u_0. There $N^2 e^{-cN} c_1^{n_1} c_2^{n_2} \le C a^k$.

Sum each bound against products of squared degree norms. Group by the index of the minimum. Its contribution is bounded by $\|(\mathcal{N}+1)\mathcal{T}_a g\|_2^2$, and the other factor by $\|g\|_\infty^2$. Orthogonality in output total degree and the triangle inequality for the two block regions give (2). The same estimates on tails give convergence as the total-degree cutoff tends to infinity, including when the unprojected multiplier is unbounded. □

13.3 Allocating the signal and controlling endpoints

Let B be the conditional expectation of $g(Z_1)g(Z_2)$ given the Gaussian linear combination with loads $qt, q(1-t)$ and independent Gaussian noise of variance $1-q$. Order the larger load first. Hermite expansion gives

$$B = P_u(\mathcal{T}_{c_1} g \otimes \mathcal{T}_{c_2} g), \qquad c_1 = r/u, \quad c_2 = s/(1-u). \tag{3}$$

The admissible u for $c_i \le 1$ form $[r, 1-s]$. For $s > 0$, minimizing $c_1 c_2 = rs/[u(1-u)]$ gives $u = \max(1/2, r)$. With that choice the condition $c_1 c_2 \le r$ is precisely (1), and both u and c_2 are at most $\max(1/2, Q) < 1$. This optimizes the allocation used in (3).

Set $L = \mathcal{T}_{qt} g + \mathcal{T}_{q(1-t)} g$, $X = \|\mathcal{N} L\|_2$. On a compact positive weight interval, freeze the allocation in (3) at the evaluation point. Differentiating the true Hermite weight factors twice costs at most $C(n_1+n_2)^2$ per coefficient. The inverse output number divides by n_1+n_2 on positive degrees, and is set to zero on degree zero. Consequently Lemma 13.2, applied with $a = r$, gives

$$\|\mathcal{N} B\|_2 + \|\mathcal{N}^{-1} B''\|_2 \le C \|g\|_\infty X, \tag{4}$$

since $\|(\mathcal{N}+1)\mathcal{T}_r g\|_2 \le CX$. All active degrees of g are even and at least four.

For $q \in [1/2, Q]$ the margin $r - q^2 = q(1-q) - s$ is uniformly positive on a sufficiently small common endpoint neighborhood. The complex weight factors are polynomials in $z, 1-z$, since the individual input degrees are even. On a small complex disk around any such real weight, the absolute total load $q(|z| + |1-z|)$ lies uniformly below a level b' with $(b')^2 < r$ at the real center. Remove complex phases by unitary input degree multipliers and use conditional contraction for the absolute loads. The gap to b' pays one output number factor. Spectral Hölder gives $\|\mathcal{T}_{b'} g\|_2^2 \le \|g\|_\infty \|\mathcal{T}_r g\|_2$ because $(b')^2 < r$. The resulting analytic L^2 bounds for $\mathcal{N} B$ and Cauchy's estimate give (4) uniformly there. The same spectral margin allows the Hermite series to converge analytically in L^2.

It suffices to treat $q \ge 1/2$, increasing Q to $1/2$ if necessary. At a smaller signal apply the result at $1/2$ to $\mathcal{T}_{2q} g$, which preserves admissibility, contracts the sup norm and is injective.

13.4 Entropy curvature

Let $R_0 = L+B$ be the Gaussian-relative output correction. The covariance is I_d, so the nonconstant part of entropy is $-\int(1+R_0)\log(1+R_0)\,d\gamma_d$. Conditioning bounds $\|R_0\|_\infty$ by $(1+\|g\|_\infty)^2-1$. For g sufficiently small the density ratio is bounded away from zero. Write $\psi(z)=\log(1+z)-z$. Differentiation gives

$$F'' = -\langle L''+B'', L+B+\psi(R_0)\rangle - \int \frac{(R_0')^2}{1+R_0}\,d\gamma_d. \tag{5}$$

The L^2-analytic weight derivatives at the endpoints and the uniform positive bounded ratio give the continuous one-sided extension of (5).

In degree $2k$, the ratio of the multipliers of L'' and L is between $k(k-1)$ and $8k(k-1)$, using $\max(t,1-t)\ge 1/2$. Consequently $\langle L'', L\rangle \ge cX^2$ and $\|\mathcal{N}^{-1}L''\|_2 \le CX$. For a smooth bounded Gaussian function f, integrating the Gaussian divergence of $f|\nabla f|^2\nabla f$ proves

$$\||\nabla f|^2\|_2 \le 3\|f\|_\infty \|\mathcal{N}f\|_2,$$

using $\|\nabla^2 f\|_{L^2(\mathrm{HS})} \le \|\mathcal{N}f\|_2$ and Cauchy–Schwarz. The Gaussian number chain rule thus gives $\|\mathcal{N}\psi(R_0)\|_2 \le C\|R_0\|_\infty\|\mathcal{N}R_0\|_2$. Approximation by Gaussian smoothing gives the same bound on bounded functions in the number domain. Pair (5) through $\mathcal{N}^{-1}$ and insert (4). Apart from the nonpositive last integral, the errors to $-\langle L'', L\rangle$ have absolute value at most $C_Q\|g\|_\infty X^2$. This gives strict negativity for the stated radius, since $X>0$ for nonzero g. The condition (1) characterizes the allocations used here.

13.5 Endpoint curvature without smoothing

Theorem 13.3. *Let $0<q\le 1$ and let g be an admissible perturbation as defined above, with $\|g\|_\infty \le 1/10$. If $q=1$, assume $\mathcal{N}g \in L^2(\gamma_d)$. Then F'' extends continuously from the interior to both endpoints, and*

$$F''_{q,g}(0^+) = F''_{q,g}(1^-) \le -\tfrac{1}{16}\|\mathcal{N}\mathcal{T}_q g\|_2^2.$$

The scalar endpoint identity is due to Abu-Ajamieh [2, Theorem 3.1]. Section A proves a whole-class scalar estimate for log-concave inputs at small signal. Here the Gaussian-relative norm controls the dimension-independent endpoint sign all the way to the unthermalized signal.

Proof. It suffices to consider $q=1$ with $\mathcal{N}g \in L^2$, since $\mathcal{T}_q g$ satisfies this domain condition for $q<1$, with its sup norm no larger than that of g. Write $B(t)=P_{1-t}(g\otimes g)$ and $L(t)=\mathcal{T}_{1-t}g+\mathcal{T}_t g$. Here the smaller-weight factor has weight t. We justify the entropy Hessian limit at $t=0$ in the number domain.

For output degree N, the squared norm of the restriction of an adjoint projection to the blocks of smaller-weight degree J is their binomial mass as in Lemma 13.2. The same argument for derivatives inserts squares of the corresponding derivative factors. The first and second factors of $(1-t)^{(N-J)/2}t^{J/2}$, after dividing by that function, are

$$\frac{J-\mu}{2t(1-t)}, \qquad \frac{(J-\mu)^2-2(1-2t)J-2t^2N}{4t^2(1-t)^2}, \qquad \mu = Nt. \tag{6}$$

Only $J\ge 4$ occurs for $g\otimes g$. For $0<t\le 1/2$ and $\mu\le 1$, the squared numerators on $J\ge 4$ are bounded by $CJ^4\mathbf{1}_{J\ge 4} \le C(J)_4$, whose binomial expectation is at most $C\mu^4$. For $\mu\ge 1$, the first numerator has second moment at most μ, and the second has second moment at most $C\mu^2$ by the

binomial fourth central moment. Consequently the first-derivative operator on those blocks has squared norm at most CN^2, and the second-derivative operator divided by N has the same bound.

For interior t, the central-moment bounds give norm at most $C_t N$ for both derivatives without division by N, uniformly on compact interior weight intervals. Since the tensor is in the total number domain, the derivatives exist in L^2 there. By the uniform estimates and dominated convergence in total degree, $\mathcal{N}B(t)$ and $B'(t)$ tend strongly to zero at 0. The inverse-number second derivative $\mathcal{N}^{-1}B''(t)$ converges strongly to its degree-four input term. The corresponding linear limits follow from $(1-t)^k + t^k$ in degree $n = 2k \geq 4$.

We have $R_0(t) \to g$ strongly in the number domain. The densities are uniformly bounded above and away from zero by conditioning. The number of $\log(1 + R_0(t))$ converges strongly to the number of $\log(1+g)$. Indeed the gradient inequality in Section 13.4 applied to $R_0(t) - g$ gives strong L^4 convergence of gradients. The chain factors converge in measure and stay bounded, while the number and squared-gradient terms converge in L^2. Pair (5) through inverse number and use strong convergence of $R_0'(t)$ to pass to the endpoint, including its squared term. The continuous limit of the first derivative and the mean value theorem identify this with the one-sided second derivative. The same argument with $\mathcal{T}_q g \to g$ in the number domain also gives convergence of the endpoint formula as $q \uparrow 1$ when $\mathcal{N}g \in L^2$.

Put $X = \|\mathcal{N}g\|_2$ and $\varepsilon = \|g\|_\infty$. The limiting linear second-derivative multiplier is $k(k-1) + 2\mathbf{1}_{k=2}$; it gives $\langle L''(0), g\rangle \geq X^2/8$ and $\|\mathcal{N}^{-1}L''(0)\|_2 \leq X/4$. Also $R_0'(0) = -\mathcal{N}g/2$. At total degree $N = n+4$, the only surviving bilinear second-derivative block has norm $2\sqrt{\binom{n+4}{4}}$ on the two input components of degrees $n, 4$. This follows by the multinomial sum in Lemma 13.2 and differentiation of t^2. Since $2\sqrt{\binom{n+4}{4}}/(n+4) < n$ for $n \geq 4$,

$$\|\mathcal{N}^{-1}B''(0)\|_2 \leq \|g_4\|_2 X \leq \varepsilon X.$$

The chain estimate gives $\|\mathcal{N}\psi(g)\|_2 \leq \big((1-\varepsilon)^{-1} + 3(1-\varepsilon)^{-2}\big)\varepsilon X \leq 5\varepsilon X$. Thus (5) at the endpoint gives

$$F''(0^+) \leq \left[-\tfrac{1}{8} - \frac{1}{4(1+\varepsilon)} + \varepsilon + 5\varepsilon(1/4+\varepsilon)\right] X^2 \leq -\tfrac{1}{16}X^2.$$

Symmetry exchanges the endpoints. □

13.6 A spectral regularity condition for every weight

Retain $q = 1$, $\mathcal{N}g \in L^2$ and $\varepsilon \leq 1/10$. Let $X_0 = \|\mathcal{N}g\|_2$, $X(t) = \|\mathcal{N}L(t)\|_2$. The bound from (6) gives $\|\mathcal{N}^{-1}B''(t)\|_2 \leq C\varepsilon X_0$ on the left half of the weight interval. Symmetry and the endpoint limits give it everywhere. Conditional contraction of total number also gives $\|\mathcal{N}B(t)\|_2 \leq 2\varepsilon X_0$. Using these two estimates in (5), the same chain rule and the linear coercivity yield

$$F''(t) \leq -aX(t)^2 + C\big(\varepsilon X_0 X(t) + \varepsilon^2 X_0^2\big) \tag{7}$$

with positive absolute constants a, C. For instance $\|\mathcal{N}\psi(R_0)\|_2 \leq C\varepsilon(X(t) + \varepsilon X_0)$, and the linear inverse-number second derivative is bounded by $CX(t)$. Thus there are absolute $c, c_0 > 0$ such that $\varepsilon X_0 \leq cX(t)$ implies $F''(t) \leq -c_0 X(t)^2$.

If $g \neq 0$ and $\mathcal{N}^{3/2}g \in L^2$, put

$$M = \frac{\|\mathcal{N}^{3/2}g\|_2^2}{\|\mathcal{N}g\|_2^2}.$$

The squared linear multiplier is at least w^n in degree n, for $w = \max(t, 1-t)$. Under the spectral probability $n^2\|g_n\|_2^2/X_0^2$, the mean of n is M. Jensen gives $X(t)^2/X_0^2 \geq w^M$. Hence $\varepsilon \leq c2^{-M/2}$

gives strict all-weight curvature. A sufficiently small absolute radius gives strict curvature on $\min(t, 1-t) \le 1/M$, since $M \ge 4$ and $(1-1/M)^{M/2}$ is bounded below. At positive signal below one, apply the same results to the smoothed perturbation and its spectral mean. These sufficient conditions quantify the role of spectral regularity at higher signal.

14 Uniform entropy concavity for symmetric log-concave laws

14.1 Statement

Write p_q for the density of $X_q = \sqrt{q}\,X + \sqrt{1-q}\,G$, with G standard normal independent of X. The standard normal density is ϕ. For independent copies U, V of X_q set

$$F_X(q,t) = h(\sqrt{1-t}\,U + \sqrt{t}\,V),$$

where h is differential entropy. The existence clause of Luo's Problem 10.1 [7] asks for one positive signal at which the function is concave on the entire weight interval for every symmetric log-concave law. The exact endpoint bounds in Section A and the local Gaussian-relative results of Sections 2–5, 9–12.3 and [5] address related restoration regimes. Abu-Ajamieh [2] also proves an all-weight bound for a one-parameter fourth-Hermite family. We prove a common signal range on the full weight interval for the whole symmetric scalar log-concave class.

Theorem 14.1. *There exist universal $q_0, c_0 > 0$ such that for every symmetric variance-one log-concave X, every $0 < q \le q_0$, and every $t \in [0,1]$,*

$$F_{X,tt}(q,t) \le -c_0 \int \big((\log p_Z)''(z) + 1\big)^2 p_Z(z)\,dz, \qquad Z = \sqrt{1-t}\,U + \sqrt{t}\,V.$$

At the endpoints the second derivative is understood by continuous one-sided extension. The right-hand side is strictly negative unless X is Gaussian, in which case $F_X(q,t)$ is constant in t.

14.2 Fisher estimates after Gaussian smoothing

For a smooth density p write $\rho = (\log p)'$ and $s = \rho + x$ for its ordinary and relative scores. Define $I(p) = \int s^2 p$, $A(p) = \int (s')^2 p$, and $B(p) = \int (s'')^2 p$. For $p = p_q$ let $\kappa = q/(1-q)$. Differentiating the Gaussian kernel gives

$$s_q' = -\kappa + f_q, \qquad f_q(x) = \frac{q}{(1-q)^2}\,\mathrm{Var}(X \mid X_q = x), \qquad -\kappa \le s_q' \le 1.$$

The upper bound uses preservation of log-concavity under convolution. The conditional law of X is log-concave, also when the support is a bounded interval. The centered moment comparison [9], applied to that law, yields

$$|f_q'(x)| \le C f_q(x)^{3/2}.$$

Indeed f_q' is $q^{3/2}/(1-q)^3$ times its third centered moment. Thus $f_q^{-1/2}$ is uniformly Lipschitz.

Lemma 14.2. *For every centered variance-one log-concave X, uniformly for $q > 0$ sufficiently small,*

$$A(p_q) \le Cq^2, \qquad \int |s_q'|^r p_q \le C_r q^{r-2} A(p_q) \quad (r > 2).$$

Proof. The same moment comparison gives $\mathbb{E}e^{\eta|X|} \le C$ for a universal $\eta > 0$. Choose $x_0 = c_0/\sqrt{q}$ with $c_0 > 0$ small. The posterior tilt is $\exp(\sqrt{q}\,xX/(1-q) - qX^2/(2(1-q)))$. For $|x| \le x_0$, its normalizing expectation is at least $e^{-q/(2(1-q))}$ by Jensen. Its fixed absolute moments are bounded by the exponential moment strip. In particular, throughout this region

$$|s_q'| \le Cq, \qquad \rho(x) = \frac{\sqrt{q}\,\mathbb{E}[X \mid X_q = x] - x}{1-q}.$$

At $x = x_0/2$, this score is at most $-c/\sqrt{q}$; at $-x_0/2$ it is at least $c/\sqrt{q}$. Its monotonicity extends these inequalities outwards. Also $\Pr(|X_q| > x_0) \le Ce^{-c/q}$ by the latent exponential moment and the Gaussian tail. Together with the global cap on s_q' this proves $A(p_q) \le Cq^2$.

Choose K larger than the central bound for s_q'/q, and shift points where $s_q' \ge Kq$ inwards by $\delta = c_1/\sqrt{q}$. Choose K large and c_1 small with $\delta \le x_0/4$. At the shifted point y the Lipschitz bound gives

$$f_q(y) \ge q\big(K^{-1/2} + Cc_1\big)^{-2} \ge 4q.$$

In particular $|s_q'(y)| \ge q$. The above score bounds give $p_q(x) \le e^{-c'/q}p_q(y)$ along either shift. The shift is injective on each tail and their images are disjoint. Hence the entire set $s_q' \ge Kq$ has mass at most $q^{-2}e^{-c'/q}A(p_q)$. On its complement, $|s_q'| \le (K+2)q$. These estimates yield the claimed weighted bound. □

Theorem 14.3. *There is a universal $C_0 > 0$ such that for every non-Gaussian centered variance-one log-concave X,*

$$\log(I(p_{e^{-2u}})/e^{-2u})$$

is a convex function of $\Theta(u)$ at all sufficiently large u, uniformly over the class, where $\Theta'(u) = \exp(-C_0e^{-2u})$. The additive constant of Θ may be chosen so that $\Theta(u) = u + O(e^{-2u})$ as $u \to \infty$.

Proof. Put $q = e^{-2u}$, and evaluate all integrals at p_q. The Ornstein–Uhlenbeck equation is $\partial_u p = p'' + (xp)'$. Its score calculus gives

$$\dot I = -2I - 2A, \qquad \dot A = -4A - 2B + 4\int (s')^3 p.$$

For completeness, $\partial_u \log(p/\phi) = s' - xs + s^2$ and $\partial_u s = s'' + (2s - x)s' - s$. Differentiate the integrals for I, A, substitute these identities and integrate by parts against p. For $L_p h = h'' + \rho h'$, the symmetric weighted Bochner identity reads

$$\|L_p s\|^2_{L^2(p)} = B + A - \int (s')^3 p.$$

Since $\langle s, L_p s\rangle_p = -A$, Cauchy–Schwarz gives $A^2 \le I(B + A - \int (s')^3 p)$. Thus for $J = \log(I/q)$,

$$J' = -2A/I, \qquad J'' \ge -4I^{-1}\int (s')^3 p \ge 2C_0 q J'.$$

The last inequality is Lemma 14.2. As $\Theta''/\Theta' = 2C_0 q$, this proves convexity. If $I = 0$ at a positive signal, the smoothed law is Gaussian and characteristic functions give Gaussian X.

All integrations above can first be done with smooth cutoffs. On compact intervals of positive signals below one, the relative score grows at most linearly, by the two-sided Hessian cap for $\log p$. The fixed higher spatial derivatives are bounded by the centered posterior cumulant formulas, since the conditional law is strongly log-concave with variance at most $(1-q)/q$ and has centered moments bounded in terms of its variance. The time derivatives used here then have polynomial bounds by the displayed equation. The exponential moment and these bounds justify removing the cutoffs and differentiating the integrals. □

Lemma 14.4. *Let $p = p_q$ arise from symmetric log-concave X with variance one. For odd $h \in W^{1,2}(p)$ with $\mathbb{E}[X_q h(X_q)] = 0$,*

$$\mathbb{E}h(X_q)^2 \le \left(\frac{1-q}{3} + Cq\left(1 + \frac{2q}{3(1-q)}\right)\right)\mathbb{E}h'(X_q)^2.$$

Proof. Let $b_j(X) = \mathbb{E}_G[h(X_q)e_j(G)]$ for the normalized Gaussian Hermite polynomials. Conditionally on X, the Gaussian energy $(1-q)\mathbb{E}[h'(X_q)^2 \mid X]$ is the sum of $jb_j(X)^2$. Peel off one third of this energy. The remainder in the putative bound is at most $\mathbb{E}[b_0^2 + (2/3)b_1^2 + (1/3)b_2^2]$. Parity gives $\mathbb{E}b_0 = \mathbb{E}b_2 = 0$, and orthogonality to X_q gives $(\mathbb{E}b_1)^2 \le q\,\mathrm{Var}(b_0)/(1-q)$. The one-dimensional log-concave Poincaré inequality [3], applied in X, yields

$$\mathrm{Var}(b_j) \le Cq\,\mathbb{E}|\mathbb{E}_G[h'(X_q)e_j(G)]|^2.$$

The sum over $j = 0, 1, 2$ is bounded by $Cq\,\mathbb{E}h'^2$ by Bessel. The claimed coefficient follows by including the extra mean contribution for b_1. Smooth approximation gives the Sobolev version. □

14.3 A Hellinger comparison at every weight

For densities p, r put $\mathsf{H}(p,r) = \|\sqrt{p} - \sqrt{r}\|_{L^2(dx)}$. We first record a comparison that also applies to unbounded density ratios.

Lemma 14.5. *Let T be a Gaussian Ornstein–Uhlenbeck operator, including complete smoothing and the identity. For every Gaussian-relative probability density r,*

$$\|\sqrt{r} - 1\|_{L^2(\gamma)} \le C\left\|\sqrt{(r+Tr)/2} - 1\right\|_{L^2(\gamma)}.$$

Proof. Put $P = (r+Tr)/2$, $u = \min(r, 8) - 1$, and $v = (r-8)_+$. Since $P \ge r/2$, $\int v\,d\gamma \le C\int(\sqrt{P}-1)^2\,d\gamma$. Set $e = (v+Tv)/2$, so $\int e\,d\gamma = \int v\,d\gamma$ and $u + Tu = 2(P - 1 - e)$. The last expression is bounded because u is bounded and T is Markov. Where $e \le 1$, P is bounded, and $(u+Tu)^2$ is at most $C((\sqrt{P}-1)^2 + e)$. Where $e > 1$, boundedness gives the same conclusion. Thus $\|u + Tu\|_2^2 \le C\|\sqrt{P} - 1\|_2^2$. The OU spectrum is nonnegative on $L^2(\gamma)$, so $\|u\|_2 \le \|u + Tu\|_2$. Finally, pointwise $(\sqrt{r} - 1)^2 \le C(u^2 + v)$. □

Theorem 14.6. *Let X be centered variance-one and log-concave. For independent copies U, V of X_q and $Z = \sqrt{1-t}\,U + \sqrt{t}\,V$, there are universal $C, \eta > 0$ such that, for all sufficiently small $q > 0$,*

$$I(p_q) \le Cq^\eta \mathsf{H}(p_Z, \phi) \le Cq^\eta\sqrt{I(p_Z)} \qquad (0 \le t \le 1).$$

Proof. Write $a = \sqrt{1-t}$, $b = \sqrt{t}$. Let $L_a = p_{qa^2}$ and $L_b = p_{qb^2}$, including $p_0 = \phi$. They are the densities when respectively the second and the first copy in Z are replaced by an independent standard normal. Set $L = (L_a + L_b)/2$. Let $\mu = p_q$, $R = (\mu + \phi)/2$, and $R_* = (p_Z + 2L + \phi)/4$, the pushforward of $R \otimes R$ under $(x, y) \mapsto ax + by$. The signed density $D = p_Z - 2L + \phi$ is the pushforward of $(\mu - \phi) \otimes (\mu - \phi)$. Conditional Jensen with respect to $R \otimes R$ gives

$$\left(\int \frac{D^2}{R_*}\right)^{1/2} \le \int \frac{(\mu-\phi)^2}{R} \le C\mathsf{H}(\mu,\phi)^2 \le CI(\mu).$$

The last inequality follows from the Gaussian logarithmic Sobolev inequality and $\mathsf{H}^2 \le D_{\mathrm{KL}}$. Since $D = 2((p_Z+\phi)/2 - L)$, the denominator $R_* = (L + (p_Z+\phi)/2)/2$ also gives $\mathsf{H}(L, (p_Z+\phi)/2) \le CI(\mu)$.

By the triangle inequality and convexity of squared Hellinger distance, $\mathsf{H}(L,\phi) \le C(\mathsf{H}(p_Z,\phi) + I(\mu))$. If $w_* = \max(a^2,b^2)$, the weaker single-copy law is an OU smoothing of p_{qw_*}. Consequently Lemma 14.5 implies

$$\mathsf{H}(p_{qw_*},\phi) \le C\big(\mathsf{H}(p_Z,\phi) + I(p_q)\big). \tag{82}$$

To use the Fisher convexity, first note that for every fixed $0<\lambda<1$ the bound

$$D_{\mathrm{KL}}(p_s\|\phi) \le C_\lambda \mathsf{H}(p_s,\phi)^{2\lambda} \qquad (0<s\le 1/2)$$

holds uniformly. The centered entropy integrand of the density ratio, $\phi(r\log r - r + 1)$ with $r = p_s/\phi$, is at most $C(1+x^2)(\sqrt{p_s}-\sqrt{\phi})^2$, using $p_s \le C$ by Gaussian smoothing. Factor the Hellinger integrand into powers λ and $1-\lambda$ and apply Hölder. The remaining polynomial moment is bounded uniformly. Writing $s_* = qw_*$ and $s_- = (3/4)s_*$, OU entropy dissipation and Fisher monotonicity give

$$I(p_{s_-}) \le C D_{\mathrm{KL}}(p_{s_*}\|\phi).$$

Indeed the change in OU time is the fixed $\frac{1}{2}\log(4/3)$ and the relative entropy has derivative $-I$.

Use Theorem 14.3 at signals $s_-, q, 16q$. The interpolation coefficient β on the lowest signal converges, uniformly over $w_* \in [1/2,1]$, to $\log(16)/\log(16q/s_-)$ as $q \downarrow 0$. Hence β is bounded above by a constant below one and below by a constant above $1/2$. The signal prefactors in the convex bound on $\log(I/q)$ are uniformly bounded after cancellation. By Lemma 14.2 and Poincaré, $I(p_{16q}) \le Cq^2$, since the relative score has mean zero. We obtain

$$I(p_q) \le C q^{2(1-\beta)} \mathsf{H}(p_{s_*},\phi)^{2\lambda\beta}.$$

Choose $\lambda < 1$ so that $2\lambda\beta > 1$ uniformly. Since H is bounded, the right-hand side is at most $Cq^\eta \mathsf{H}(p_{s_*},\phi)$ for a fixed $\eta>0$. Absorb the $I(p_q)$ term using (82) and reduce the common signal range. The assertion is immediate for Gaussian X. □

14.4 Entropy under rotation

Let $t = \sin^2\theta$, $a = \cos\theta$, $b = \sin\theta$, $Z = aU + bV$, and $W = -bU + aV$. Use s for the relative score of U, ρ for the ordinary score of Z, and $S = \rho + z$. Put $m(z) = \mathbb{E}[W \mid Z = z]$, $D(z) = \mathbb{E}[W^2 - 1 \mid Z = z]$, $Q(z) = \mathbb{E}[s(U)s(V) \mid Z = z]$, and $c = (a^2-b^2)/(ab)$. For these definitions two exact identities hold:

$$D = S' + cm' + S^2 + cSm - Q/(ab), \tag{83}$$

$$4a^2b^2 F_{X,tt} = I(p_Z) - \mathbb{E}[S'(Z)D(Z)] - \mathbb{E}(m'(Z) + \rho(Z)m(Z))^2 + c\,\mathbb{E}[S(Z)m(Z)]. \tag{84}$$

To obtain the first identity use the two relative coordinates of the joint score, $A_0 = as(U) + bs(V)$ and $B_0 = -bs(U) + as(V)$. They satisfy $\partial_w B_0 - \partial_z A_0 = c\,\partial_z B_0$. Conditional integration by parts in w gives $D = \mathbb{E}[\partial_w B_0 + B_0^2 \mid Z]$, and the conditional means of A_0, B_0 are S, m. Differentiating these conditional means in z gives the corrections $\mathrm{Var}(A_0 \mid Z)$ and $\mathrm{Cov}(A_0, B_0 \mid Z)$, respectively. The remaining quadratic has the pointwise identity $B_0^2 - A_0^2 - cA_0B_0 = -s(U)s(V)/(ab)$. These facts give (83).

For the second identity, the density's angle derivative is $-\partial_z(p_Z m)$, and its second derivative is $\partial_z^2(p_Z\mathbb{E}[W^2 \mid Z]) + \partial_z(zp_Z)$, since $Z_\theta = W$, $Z_{\theta\theta} = -Z$. Differentiate entropy using these expressions and integrate by parts. The second angle derivative equals $\mathbb{E}[-\rho'(Z)W^2] - 1 - \mathbb{E}(m' + \rho m)^2$ and the first is $-\mathbb{E}[Sm]$, since $\mathbb{E}[ZW] = 0$. Also $\mathbb{E}S' = -I(p_Z)$. Applying the chain rule with $t_\theta = 2ab$ and $t_{\theta\theta} = 2(a^2-b^2)$ gives (84).

For regularity, U and V each include independent nondegenerate Gaussian noises, so the rotated pair has Gaussian smoothing in both coordinates. Its density derivatives have conditional

polynomial representations, and the input has exponential moments. One may differentiate angle on a compact interval in the interior and use cutoffs in the integrations. For example S' is bounded, $m = \mathbb{E}[B_0 \mid Z]$ is in every fixed $L^p(p_Z)$, and $m' = \mathbb{E}[\partial_z B_0 \mid Z] + \mathrm{Cov}(A_0, B_0 \mid Z)$ is in $L^2(p_Z)$. These statements follow from bounded input score derivatives, linear score growth, and conditional Jensen. The same estimates for the angle derivatives of p_Z, together with a Gaussian lower bound $p_Z(z) \geq c_1 e^{-c_2 z^2}$ on compact angle intervals, justify passing to the uncut entropy formulas.

14.4.1 Central weights

Fix any closed subinterval of $|c| < 2$. The latent variable $aX_1 + bX_2$ is symmetric, log-concave and variance-one. Thus Lemma 14.4 applies to p_Z. Both S and m are odd and orthogonal in $L^2(p_Z)$ to z. This follows for S from integration by parts and variance one, and for m from $\mathbb{E}[ZW] = 0$. On this subspace write

$$\mathcal{E}(f, g) = \mathbb{E}[f'g' - fg].$$

For small q, $\mathcal{E}(f, f)$ is bounded below by a fixed positive multiple of $\mathbb{E}(f')^2$, by Lemma 14.4. The square in (84) has the exact identity

$$\mathbb{E}(m' + \rho m)^2 = \mathcal{E}(m, m) + \mathbb{E}[(2 - S')m^2].$$

Since $-\kappa \leq S' \leq 1$, the cubic terms satisfy

$$\begin{aligned} -\mathbb{E}[S'(S^2 + cSm)] &= -\mathbb{E}[S'(S + (c/2)m)^2] + (c^2/4)\mathbb{E}[S'm^2] \\ &\leq \kappa \mathbb{E}(S + (c/2)m)^2 + (c^2/4)\mathbb{E}m^2. \end{aligned}$$

On $[\tau, 1 - \tau]$ we have $|c| \leq 2 - \delta_\tau$ for some $\delta_\tau > 0$. The square's excess $\mathbb{E}[(2 - S')m^2] \geq \mathbb{E}m^2$ absorbs the second term. The remaining quadratic form $\mathcal{E}(S, S) + c\mathcal{E}(S, m) + \mathcal{E}(m, m)$ controls $(1 - |c|/2)\big(\mathcal{E}(S, S) + \mathcal{E}(m, m)\big)$. The term with κ is absorbed for small q.

Finally $\|Q\|_{L^2(p_Z)} \leq I(p_q)$ by conditional Jensen and independence. By Theorem 14.6 and Lemma 14.4 gives this is $O(q^\eta \sqrt{A(p_Z)})$. Hence the Q-term in (84) is $O_\tau(q^\eta A(p_Z))$, which is also absorbed. Since $4a^2b^2 \leq 1$, this proves the claimed estimate on this subinterval. In particular it covers $1/5 \leq t \leq 4/5$.

14.5 A linear Fisher comparison with a buffer

Let $\omega(x) = 1 + x^2$, and let H_ω denote the Hellinger norm using this weight. We will use the following strengthened form of the argument from Section 3:

$$I(p_{q/2}) \leq C A(p_Z) \qquad \text{if } \max(a^2, b^2) \geq 4/5. \tag{85}$$

We give details because the linear control by $A(p_Z)$ will be used for the weak copy at all exterior weights.

For a Gaussian-relative density r, Gaussian integration by parts and logarithmic Sobolev give $\mathsf{H}_\omega(r\phi, \phi)^2 \leq C I(r\phi)$. Indeed for $f = \sqrt{r} - 1$, $\|xf\|_2 \leq C(\|f\|_2 + \|f'\|_2)$ follows by integrating $\int x^2 f^2 \, d\gamma$ once, and $4\|f'\|_2^2 = I(r\phi)$. The clipping argument in Lemma 14.5 also applies with the weight ω. The identity for $u + Tu$ and its pointwise clipping bound are unchanged, and $\int \omega T v \, d\gamma \leq C \int \omega v \, d\gamma$. To conclude, $(I + T)^{-1}$ is uniformly bounded on $L^2(\omega \, d\gamma)$. If T has Hermite multipliers r_0^n, the inverse has multipliers $m_n = (1 + r_0^n)^{-1}$. In the Hermite basis multiplication by x connects degrees $n, n + 1$ with coefficient $\sqrt{n + 1}$. Thus its commutator with the inverse is bounded on unweighted L^2, since $\sup_{n \geq 0} \sqrt{n + 1}\, |m_{n+1} - m_n|$ is bounded uniformly in $r_0 \in [0, 1]$.

This bound follows from $|m_{n+1} - m_n| \le (1 - r_0) r_0^n$, with the constant Hermite degree multiplier of T set to one. Consequently $\mathsf{H}_\omega(p_{qw_*}, \phi) \le C\mathsf{H}_\omega(L, \phi)$.

Repeat the product calculation with the output weight $\omega(z)$. The inequality $\omega(ax + by) \le C(\omega(x) + \omega(y))$, followed by conditional Jensen, gives

$$\left(\int \omega \frac{(p_Z - 2L + \phi)^2}{R_*} \right)^{1/2} \le C\mathsf{H}_\omega(p_q, \phi)\mathsf{H}(p_q, \phi) \le CI(p_q).$$

The weighted Hellinger triangle inequality with Theorem 14.6, and Lemma 14.4 gives $\mathsf{H}_\omega(p_{qw_*}, \phi) \le C\sqrt{A(p_Z)}$.

For any fixed compact range of OU correlation $\lambda \in (0, 1)$, the corresponding OU operator T_λ satisfies

$$I((T_\lambda r)\phi) \le C\mathsf{H}_\omega(r\phi, \phi)^2. \tag{86}$$

To see this, let $Y = \lambda x + \sqrt{1 - \lambda^2}\, G$ and $M(x) = \mathbb{E}r(Y)$. A Gaussian kernel derivative makes $M'(x)$ a fixed bounded multiple of $\mathbb{E}[Gr(Y)]$. Where $M \ge 1/2$, write $r - 1 = (\sqrt{r} - 1)(\sqrt{r} + 1)$ and use Cauchy–Schwarz to bound $|M'|^2/M$ by $C\mathbb{E}[G^2(\sqrt{r(Y)} - 1)^2]$. Where $M < 1/2$, conditional Cauchy–Schwarz instead bounds it by $C\mathbb{E}[G^2 r(Y)]$. This is at most $C + C\mathbb{E}[G^2(\sqrt{r(Y)} - 1)^2]$, and the constant is absorbed by $C\mathbb{E}[(\sqrt{r(Y)} - 1)^2]$, since $\mathbb{E}\sqrt{r(Y)} \le \sqrt{M}$. Integrating over $x \sim \gamma$, the conditional second moment of G given Y is at most $1 + Y^2$. This proves (86). Taking $\lambda^2 = 1/(2w_*)$ proves (85), since these correlations are in a fixed compact interval.

14.6 The weaker copy

Consider $a \ge b > 0$ with $b^2 \le 1/5$. Set $v = q/2$, $d = \sqrt{2}b$, and $\sigma_1 = 1/2$. With independent X_1, Y, G_1, G_2 and $Y \sim p_v$, put

$$R = dY + \sigma_1 G_1, \qquad Z = a\sqrt{q}X_1 + R + \sigma_2 G_2, \qquad \sigma_2^2 = 1 - qa^2 - d^2 - \sigma_1^2.$$

This gives the original output law and both noise variances are bounded below. If $J = \mathbb{E}[s_q(V) \mid aU + bV]$, then in this representation

$$J(z) = \sqrt{2}\, \mathbb{E}[s_v(Y) \mid Z = z]. \tag{87}$$

For the Fourier convention $+i\xi x$, integration by parts gives

$$\widehat{s_u p_u}(\xi) = \sqrt{u}\, e^{-(1-u)\xi^2/2}[-i\chi_X'(\sqrt{u}\,\xi) - i\sqrt{u}\,\xi\,\chi_X(\sqrt{u}\,\xi)],$$

where χ_X is the characteristic function of X. Apply this formula at $(q, b\xi)$ and $(v, d\xi)$. The Gaussian factors agree after including all the remaining noise and the independent strong arm. The same argument permits separate allocations for the two factors of a signed score product.

Write $e_j = \mathrm{He}_j/\sqrt{j!}$ and $A_j(u) = \mathbb{E}[s_u(X_u)e_j(X_u)]$. Integration by parts and Hermite smoothing give

$$A_j(u) = \sqrt{j+1}\, \mathbb{E}e_{j+1}(X_u) = (u/v)^{(j+1)/2} A_j(v).$$

These moments vanish for even j and for $j = 1$.

Lemma 14.7. *Let $\|h\|_{1,Z} = (\mathbb{E}h(Z)^2 + \mathbb{E}h'(Z)^2)^{1/2}$. For every fixed $b_* > 0$ and $\epsilon > 0$, there are $N < \infty$ and a common small signal range on which*

$$\left\| J - \sum_{\substack{3 \le j \le N \\ j \text{ odd}}} b^j A_j(q) e_j \right\|_{1,Z} \le \epsilon \sqrt{I(p_v)} \quad (b_* \le b \le 1/\sqrt{5}).$$

As $(b,q)\to(0,0)$, uniformly over the input class,

$$\|J-b^3A_3(q)e_3\|_{1,Z}=o(b^3)\sqrt{I(p_v)}.$$

Proof. The first conditional map $h\mapsto\mathbb{E}[h(Y)\mid R]$ has kernel on Lebesgue L^2, after conjugation by square roots of densities,

$$K_d(x,y)=\frac{\sqrt{p_v(y)}\,\phi_{\sigma_1}(x-dy)}{\sqrt{p_R(x)}},$$

where ϕ_{σ_1} has variance σ_1^2. If the log slope of p_v is at most $-M$ on $[T-\Delta-1,\infty)$ with $T-\Delta-1>0$, bound p_R below by integrating on $[y-\Delta-1,y-\Delta]$ when $y>T$. On this interval $p_v\ge e^{M\Delta}p_v(y)$. Completing the square in the ratio of the Gaussian kernel to the square root of its shifted minimum gives

$$K_d(x,y)\le Ce^{-M\Delta/2+Cd^2(\Delta+1)^2}e^{-c(x-dy)^2}.$$

The other tail is symmetric. Schur's test, integrating the last factor in x and y, bounds the operator norm on $|y|>T$ by

$$Cd^{-1/2}e^{-M\Delta/2+Cd^2(\Delta+1)^2}. \tag{88}$$

For d in a compact subinterval of $(0,1)$, first fix large T, then take q small. The posterior bounds give normal score convergence on fixed compact intervals uniformly in X, so by log-concavity one can take $M=cT$ and $\Delta=\alpha T$. Choose the fixed $\alpha>0$ small enough to give a bound $Ce^{-c'T^2}$ in (88), also valid for normal input. On $|y|\le T$, the kernel converges on compact sets to its Gaussian counterpart uniformly over X,d, by the natural L^2 coupling $Y\to G$ and Gaussian convolution. The bound $p_R(x)\ge c\phi_{\sigma_1}(x)$ follows by restricting to $|Y|\le 2$ and averaging symmetric shifts. The ratio of the two-shift average to the centered Gaussian is $e^{-d^2Y^2/(2\sigma_1^2)}\cosh(dxY/\sigma_1^2)$. This lower bound gives an integrable squared Gaussian majorant on truncated y and proves Hilbert–Schmidt convergence there. With (88) this gives operator-norm convergence of the first map.

For normal Y put $r=(d^2+\sigma_1^2)^{1/2}$. This map sends $e_j(Y)$ to $(d/r)^je_j(R/r)$. Since $d/r<1$ uniformly, truncating at degree N gives arbitrarily small operator error. In the finite input and output basis vectors one may replace the Gaussian square-root density by the actual square-root density: convergence in Lebesgue L^2 follows from local convergence and uniform high moments.

The second map $h\mapsto\mathbb{E}[h(R)\mid Z]$ takes $L^2(p_R)$ boundedly into $\|\cdot\|_{1,Z}$. Its derivative is $-\operatorname{Cov}(h(R),G_2\mid Z)/\sigma_2$. The conditional density of G_2 is Gaussian times a log-concave factor from $R+a\sqrt{q}X_1$, so its variance is at most one [14]. Conditional Cauchy–Schwarz gives the bound on the derivative. It and conditional Jensen give the same bounds from $L^p(p_R)$ to $L^p(p_Z)$ for $p\ge 2$. Applied to $e_j(R/r)$, the second map converges in $\|\cdot\|_{1,Z}$ to $r^je_j(Z)$ as $q\to 0$, uniformly in d. On compact sets use the explicit Gaussian kernel for Z to differentiate numerator and denominator of the conditional expectation. The inputs converge in all fixed moments by coupling, and the Gaussian variances are uniformly bounded away from zero. Outside a compact set, the L^p bounds with $p>2$ and uniform moments control the tail. This convergence includes $d\to 0$ since $r\ge\sigma_1$. Compose the maps and use $\sqrt{2}\,d^jA_j(v)=b^jA_j(q)$ to prove the first claim.

For $d\downarrow 0$, take $T=d^{-1/2}$, $\Delta=T/3$. The negative log slope at a fixed positive point for small q gives a constant $M>0$ in (88). The resulting tail operator norm is smaller than every fixed power of d. On $|y|\le T$, expand $\phi_{\sigma_1}(x-dy)$ through degree three in dy. Its fourth derivative translated by $|dy|\le\sqrt{d}$ has bounded $L^2(p_R^{-1}dx)$ norm by the lower Gaussian bound. After integration against hp_v the remainder costs $Cd^4\|h\|_{L^2(p_v)}$ by the uniform eighth moment. Extending the polynomial moments to all y costs less by exponential moments. For $h=s_v$, degrees zero through two vanish and the cubic term is

$$-\frac{d^3\sqrt{6}\,A_3(v)}{6}\frac{\phi'''_{\sigma_1}}{p_R}.$$

As $d \to 0$, ϕ'''_{σ_1}/p_R converges to $\phi'''_{\sigma_1}/\phi_{\sigma_1}$ in $L^2(p_R)$. Use compact convergence, $p_R \geq c\phi_{\sigma_1}$ and uniform high moments for the limiting polynomial. Now $-\phi'''_{\sigma_1}/\phi_{\sigma_1} = \sqrt{6}\,\sigma_1^{-3} e_3(\cdot/\sigma_1)$. The second map above with $r \to \sigma_1$ gives exactly the endpoint expansion after multiplying by $\sqrt{2}$. The first-map bound is also valid with any fixed positive variance s^2 in place of σ_1^2:

$$\|\mathbb{E}[s_v(Y) \mid dY + sG]\|_2 \leq C_s d^3 \sqrt{I(p_v)}. \tag{89}$$

□

14.7 Exterior weights and the endpoints

Continue to assume $a \geq b$ and $b^2 \leq 1/5$. Write $K = A(p_Z)$, $k = b^2/a^2$. The two conditional score identities give $m = -bS/a + J/a$. Also J is odd and orthogonal to z, since $\mathbb{E}[Zs_q(V)] = b\,\mathbb{E}[Vs_q(V)] = 0$. Discarding the nonpositive excess $-\mathbb{E}[(2 - S')m^2]$, the quadratic part of (84) becomes

$$-2k\mathcal{E}(S,S) + \frac{2b/a - c}{a}\mathcal{E}(S,J) - a^{-2}\mathcal{E}(J,J). \tag{90}$$

The coefficient in the middle is nonpositive because $b^2 \leq 1/5 < 1/4$. The last term is nonpositive by the gap inequality and the first is at most $-c_1 b^2 K$. We control the middle term by the signs of the finite Hermite pairings. For any fixed N, the score identity and the Hermite addition formula give, for $j \leq N$,

$$\mathbb{E}[Se_j(Z)] = (a^{j+1} + b^{j+1})A_j(q) + O_N(K).$$

Indeed each other term in the addition formula is a product of two positive-degree Hermite moments of U, V. Each is $O_N(\sqrt{I(p_v)})$ by the score moment formula and the fixed signal ratio $q/v = 2$, and $I(p_v) \leq CK$ by (85). For $j = 3$ the formula is exact, since degrees one through three of the density moments vanish. Integrating by parts in the derivative pairing yields

$$\mathcal{E}(S, e_j) = (j-1)\mathbb{E}[Se_j(Z)] - \mathbb{E}[S^2 e_j'(Z)].$$

For a bound on the error, use Poincaré on S^2 and Cauchy–Schwarz:

$$\|S\|_4^2 \leq I(p_Z) + C\|S'\|_4^2 \leq Cq\sqrt{K}.$$

The last step uses Lemma 14.2, $I(p_Z) \leq CK$ and $K \leq Cq^2$. In particular $\|SS'\|_2 \leq Cq\sqrt{K}$, and the polynomial error above is $O_N(q\sqrt{K})$.

On any compact range $b_* \leq b \leq 1/\sqrt{5}$ use Lemma 14.7. The leading summands in the pairing $\mathcal{E}(S, J)$ are

$$(j-1)b^j(a^{j+1} + b^{j+1})A_j(q)^2 \geq 0.$$

The errors from the finite terms are $O_N(qK)$, since each $A_j(q) = O_N(\sqrt{K})$. The approximation error in the pairing is $C\epsilon K$ by (85) and the gap bound on S. It follows by first choosing ϵ, then the common small signal range, that the middle term of (90) costs arbitrarily little relative to $b^2 K$ on the compact range.

For $(b, q) \to (0, 0)$, the cubic approximation gives, by the same sign on the cubic pairing,

$$\mathcal{E}(S, J) \geq -o(b^3 K), \qquad \|J\|_{1,Z} \leq Cb^3\sqrt{K}.$$

Here the little-oh is uniform over X. Since the middle coefficient in (90) is $O(1/b)$, this yields an error $o(b^2 K)$. The nonlinear expression from (83) is $S^2 + cSm = kS^2 + (c/a)SJ$. The contribution of the first term to $-\mathbb{E}[S'(S^2 + cSm)]$ is at most $k\kappa I(p_Z)$; the second is bounded in absolute value

by $Cb^{-1}\|SS'\|_2\|J\|_2$. Thus their sum is at most Cqb^2K near the endpoint. On compact exterior weights the same estimates apply with $\|J\|_{1,Z} \le C\sqrt{K}$.

The Q-term in (84) is $O(q^\eta K)$ on compact exterior weights by Theorem 14.6. Near $b = 0$ allocate the strong score at $v_h = 5q/4$ and the weak at $v = q/2$. The two gains are $a/\sqrt{5/4}$ and $d = \sqrt{2}b$. For all small b their squared sum is bounded away from one. The signed score identity used for (87) represents Q, up to a fixed factor, by projecting $s_{v_h}(X_{v_h})s_v(Y)$ onto the output with these gains and spare Gaussian noise. Reserve a fixed positive share of the spare variance for the weak channel, and condition first on its output and the independent strong input. By (89) and conditional contraction,

$$\|Q\|_{L^2(p_Z)} \le Cb^3\sqrt{I(p_v)I(p_{v_h})} \le Cqb^3\sqrt{K}.$$

The last inequality uses (85) and Lemma 14.2 with Poincaré. Its contribution to $4a^2b^2F_{X,tt}$ is consequently $O(qb^2K)$.

Choose the endpoint window $0 < b \le b_*$ small enough and then the signal range small enough, so that the errors preserve a fixed fraction of the first term in (90). For the compact exterior interval $[b_*, 1/\sqrt{5}]$ decrease the signal range as needed by the compact argument. It follows that $4a^2b^2F_{X,tt} \le -cb^2K$ on the entire exterior region. Dividing gives $F_{X,tt} \le -c'K$. Taking the minimum of the constants from the central and exterior regions proves the asserted bound throughout the open interval.

At a weight endpoint use the even parametrization in $b = \sqrt{t}$, with the other coefficient $\sqrt{1-b^2}$. The Gaussian-kernel regularity argument in Section A shows that the entropy in this variable is even and C^4; its fourth Taylor coefficient gives the continuous one-sided t-second derivative. The quantity $A(p_Z)$ is continuous there by local Gaussian kernel convergence, the uniform Hessian cap and tail tightness. The bound therefore extends to both endpoints. Finally $A(p_Z) = 0$ forces Z Gaussian. Symmetry and exponential moments identify X from the cumulants: its odd cumulants are zero, and at every even order $j \ge 4$ its cumulant is multiplied in Z by the strictly positive $q^{j/2}(a^j + b^j)$. Thus the moment generating function of X agrees with the Gaussian near zero, and X is Gaussian.

The universal concavity signals form an initial interval. If concavity holds universally at signal q_1, apply it to the smoothed input X_{q/q_1} to obtain it at any $0 < q < q_1$. The class is preserved by this smoothing. Continuity of entropy at positive signals shows that the supremum q_* is included, and Luo's symmetric counterexamples give $q_* \le 4/7$. The theorem proves $q_* > 0$.

15 Entropy concavity for isotropic log-concave vectors

15.1 Statement

Fix $n \ge 1$. For a centered isotropic log-concave random vector $X \in \mathbb{R}^n$, let $X_q = \sqrt{q}\,X + \sqrt{1-q}\,G$, where G is independent standard normal. Let U, V be independent copies of X_q. Put $Z = \sqrt{1-t}\,U + \sqrt{t}\,V$ and $F_X(q,t) = h(Z)$, with h differential entropy.

Theorem 15.1. *There are $q_n, c_n > 0$ depending only on n such that, for every centrally symmetric isotropic log-concave X, all $0 < q \le q_n$, and all $t \in [0,1]$,*

$$F_{X,tt}(q,t) \le -c_n \int \|\nabla^2 \log p_Z(z) + I_n\|_{\mathrm{HS}}^2 p_Z(z)\,dz.$$

The weight derivatives at $0, 1$ are continuous one-sided limits. The right-hand side is strictly negative for every non-Gaussian X.

The universal concavity signals in dimension n are an interval $(0, q_{*,n}]$, with $0 < q_{*,n} \le 4/7$. The upper bound follows from Luo's symmetric scalar examples [7] by adjoining independent standard normal coordinates. The signals form a downward-closed set by smoothing the input at the ratio of two signals, and closedness at positive signals follows from continuity of entropy. The result concerns curvature on the entire weight interval. The entropy-jump results of Ball and Nguyen [15] compare the equal-weight entropy to the single-copy entropy, and the bounded Gaussian-relative result of Sections 11–11.4.2 concerns a neighborhood of the Gaussian. The scalar theorem Section 14 proves the existence clause of Luo's Problem 10.1; here we allow arbitrary dependence between the n coordinates.

15.2 Fisher interpolation and comparison

Let p_q be the density of X_q, ϕ_n the standard normal density, $s = \nabla \log(p/\phi_n)$, and $H = \nabla s$. Write

$$I(p) = \mathbb{E}|s|^2, \qquad A(p) = \mathbb{E}\|H\|_{\mathrm{HS}}^2, \qquad B(p) = \mathbb{E}\|\nabla H\|_{\mathrm{HS}}^2.$$

All norms and expectations are under p unless a different measure is specified. Constants in this paper may depend on n. Isotropic log-concave laws have a common exponential moment in fixed dimension, and a common Poincaré bound [16]. For the former claim, apply the centered one-dimensional moment comparison [9] to coordinate marginals and use Hölder. Exponential moments and the common Poincaré bound apply also to the log-concave isotropic latent sum $aX_1 + bX_2$ when $a^2 + b^2 = 1$.

Lemma 15.2. *Uniformly on the centered isotropic log-concave class in fixed dimension, for all sufficiently small $q > 0$,*

$$A(p_q) \le C_n q^2, \qquad \mathbb{E}\|H\|_{\mathrm{HS}}^r \le C_{n,r} q^{r-2} A(p_q) \quad (r > 2).$$

Also $\log(I(p_q)/q)$, with $q = e^{-2u}$, is convex in the time $\Theta(u)$ defined by $\Theta'(u) = \exp(-C_n e^{-2u})$, unless X is Gaussian.

Proof. With $\kappa = q/(1-q)$, differentiation of the Gaussian kernel gives

$$H = -\kappa I_n + M, \qquad M = \frac{q}{(1-q)^2}\,\mathrm{Cov}(X \mid X_q = x), \qquad -\kappa I_n \preceq H \preceq I_n.$$

The posterior is log-concave and nondegenerate. Let $\lambda(x) = \|M(x)\|_{\mathrm{op}}$. Directionally differentiating its quadratic forms gives posterior third centered mixed moments. The centered log-concave moment comparison and Hölder show, for unit vectors e, h,

$$|D_h(e^{\mathsf{T}} M e)| \le C\lambda^{3/2}.$$

The same bound holds for directional derivatives of λ where they exist, by the variational formula for the largest eigenvalue. Hence $\lambda^{-1/2}$ is uniformly Lipschitz on each line.

For $|x| \le r_0 = \epsilon_n/\sqrt{q}$, posterior moments are uniformly bounded. The tilt is $\exp(\sqrt{q}\,x \cdot X/(1-q) - q|X|^2/(2(1-q)))$; Jensen bounds its normalizer below by $e^{-qn/(2(1-q))}$, and the common exponential moment controls its numerator. Thus $\|H\|_{\mathrm{op}}$ is $O_n(q)$ on this ball, and the outward radial derivative of $\log p_q$ at radius $r_0/2$ is at most $-c_n/\sqrt{q}$ by the posterior mean formula. Log-concavity extends this radial slope bound outward. The probability of $|X_q| > r_0$ is $O_n(e^{-c_n/q})$ by the latent exponential and Gaussian tails. This proves $A(p_q) \le C_n q^2$.

To bound the weighted moments, on $\{\lambda \geq K_n q\}$ move radially inwards by $\delta = \epsilon_n'/\sqrt{q}$ with $0 < \epsilon_n' \leq \epsilon_n/4$. Choose K_n large and ϵ_n' small. The shifted point y has $\lambda(y) \geq q(K_n^{-1/2} + C\epsilon_n')^{-2} \geq 4q$, hence $\|H(y)\|_{\mathrm{HS}} \geq q$. The density at the original point is at most $e^{-c_n/q} p_q(y)$. The radial change of variables is injective and has Jacobian at least $(3/4)^{n-1}$. Therefore

$$\Pr(\lambda \geq K_n q) \leq C_n q^{-2} e^{-c_n/q} A(p_q).$$

On its complement $\|H\|_{\mathrm{HS}} \leq C_n q$, and globally $\|H\|_{\mathrm{HS}} \leq C_n$. This proves the first assertion.

For the OU evolution the identities are

$$\dot{I} = -2I - 2A, \qquad \dot{A} = -4A - 2B + 4\mathbb{E}\,\mathrm{tr}(H^3), \qquad \|L_p s\|_2^2 = B + A - \mathbb{E}\,\mathrm{tr}(H^3), \quad L_p = \Delta + \nabla \log p \cdot \nabla.$$

Indeed $\partial_u s = \Delta s + (2s - x) \cdot \nabla s - s$, and differentiating in space gives $\partial_u H = \Delta H + (2s - x) \cdot \nabla H + 2H^2 - 2H$. Differentiating the integrals and using the OU Fokker–Planck equation gives the first two identities. The last is the weighted Bochner formula componentwise. Cauchy–Schwarz applied to $-\langle s, L_p s\rangle = A$ now shows, for $J = \log(I/q)$,

$$J' = -2A/I, \qquad J'' \geq -4\mathbb{E}\,\mathrm{tr}(H^3)/I \geq 2C_n q J'.$$

Here $|\,\mathrm{tr}(H^3)| \leq \|H\|_{\mathrm{HS}}^3$. Since $\Theta''/\Theta' = 2C_n q$, this is convexity in Θ. The identities can be integrated first with smooth cutoffs. On compact subintervals of signals in $(0, 1)$, the posterior covariance bound, centered moment comparison and posterior derivative formulas give bounded higher spatial log derivatives, linear score growth, and polynomially bounded time derivatives. The uniform exponential tails justify the limits. □

Lemma 15.3. *For $a, b \geq 0$, $a^2 + b^2 = 1$, and $Z = aU + bV$, there are $C_n, \eta_n > 0$ such that*

$$I(p_q) \leq C_n q^{\eta_n} \mathsf{H}(p_Z, \phi_n) \leq C_n q^{\eta_n} \sqrt{I(p_Z)}$$

uniformly in the weights for all sufficiently small q. In addition, if $w_* = \max(a^2, b^2) \geq 4/5$, then

$$I(p_{q/2}) \leq C_n A(p_Z). \tag{91}$$

Proof. Write H for the $L^2(dx)$ distance of square roots. With $\mu = p_q$, put $L = (p_{qa^2} + p_{qb^2})/2$, $R = (\mu + \phi_n)/2$, and $R_* = (p_Z + 2L + \phi_n)/4$. The latter is the pushforward of $R \otimes R$ under the weighted addition map. The signed density $D = p_Z - 2L + \phi_n$ is the pushforward of $(\mu - \phi_n) \otimes (\mu - \phi_n)$. Conditional Jensen gives

$$\left\| D/\sqrt{R_*} \right\|_2 \leq \int (\mu - \phi_n)^2/R \leq C\mathsf{H}(\mu, \phi_n)^2 \leq CI(\mu),$$

using Gaussian logarithmic Sobolev. The denominator is half the sum of L and $(p_Z + \phi_n)/2$, so the Hellinger triangle inequality gives $\mathsf{H}(L, \phi_n) \leq C(\mathsf{H}(p_Z, \phi_n) + I(\mu))$.

One can recover the less smoothed density in this mixture. For $r \geq 0$ with $\int r\, d\gamma = 1$ and T a Gaussian OU operator, let $P = (r + Tr)/2$ and split $r - 1 = u + v$ with $u = \min(r, 8) - 1$, $v = (r - 8)_+$. Since $P \geq r/2$, $\int v\, d\gamma \leq C\mathsf{H}(P\phi_n, \phi_n)^2$. With $e = (v + Tv)/2$, $\int e = \int v$ under γ and $u + Tu = 2(P - 1 - e)$ is bounded. On $e \leq 1$, P is bounded, so $(u + Tu)^2 \leq C((\sqrt{P} - 1)^2 + e)$; on $e > 1$ use boundedness. The OU spectrum is nonnegative, so $\|u\|_2 \leq \|u + Tu\|_2$ and $\mathsf{H}(r\phi_n, \phi_n)^2 \leq C(\|u\|_2^2 + \int v\, d\gamma)$. Since the stronger arm p_{qw_*} and its OU smoothing make up L, this shows

$$\mathsf{H}(p_{qw_*}, \phi_n) \leq C(\mathsf{H}(p_Z, \phi_n) + I(p_q)). \tag{92}$$

For $s \le 1/2$, $p_s \le C_n$ by Gaussian smoothing. The centered KL integrand is at most $C_n(1+|x|^2)(\sqrt{p_s}-\sqrt{\phi_n})^2$. Hölder and uniform moments then give $D_{\mathrm{KL}}(p_s\|\phi_n) \le C_{n,\ell}\mathsf{H}(p_s,\phi_n)^{2\ell}$ for every fixed $0<\ell<1$. OU entropy dissipation and monotonicity give $I(p_{3s/4}) \le CD_{\mathrm{KL}}(p_s\|\phi_n)$. Interpolate $\log(I/q)$ at $3qw_*/4, q, 16q$ by Lemma 15.2. The coefficient β on the lowest signal approaches $\log(16)/\log(64/(3w_*))$ uniformly in $w_* \in [1/2,1]$, so it has a strict uniform margin above $1/2$ and below one. The signal factors cancel up to bounded constants. With $I(p_{16q}) \le C_n q^2$ by the Fisher lemma and Poincaré, this gives

$$I(p_q) \le C_{n,\ell} q^{2(1-\beta)} \mathsf{H}(p_{qw_*},\phi_n)^{2\ell\beta}.$$

Choose ℓ so that $2\ell\beta > 1$. Absorb the $I(p_q)$ term from (92) to prove the first claim. The second inequality follows from Gaussian logarithmic Sobolev. Gaussian inputs trivially satisfy both claims.

To obtain (91), repeat the clipping calculation with $\omega(x) = 1+|x|^2$. Write H_ω for the corresponding Hellinger norm. The clipping tail estimate is unchanged with the weight, since $T\omega \le C\omega$. The inverse of $I+T$ is bounded on $L^2(\omega\, d\gamma)$ uniformly in the correlation r_0. Its multipliers are $m_k = (1+r_0^k)^{-1}$ on total Hermite degree k. Commuting the inverse with each coordinate multiplication costs at most $C\sup_k \sqrt{k+1}|m_{k+1}-m_k|$, bounded uniformly by the two adjacent Hermite-degree shifts. This proves the weighted clipping estimate. Gaussian integration by parts for $f = \sqrt{r}-1$ gives $\||x|f\|_2 \le C_n(\|f\|_2 + \|\nabla f\|_2)$ and $4\|\nabla f\|_2^2 = I(r\phi_n)$. The signed product transfer with this weight costs at most $C\mathsf{H}(\mu,\phi_n)\mathsf{H}_\omega(\mu,\phi_n)$, since $\omega(ax+by) \le C(\omega(x)+\omega(y))$. Consequently

$$\mathsf{H}_\omega(p_{qw_*},\phi_n) \le C_n(\mathsf{H}_\omega(p_Z,\phi_n) + I(p_q)) \le C_n\sqrt{A(p_Z)}.$$

For a fixed compact correlation range $\lambda \in (0,1)$, $I((T_\lambda r)\phi_n) \le C_n \mathsf{H}_\omega(r\phi_n,\phi_n)^2$. Indeed with $Y = \lambda x + \sqrt{1-\lambda^2}\,G$ and $M = \mathbb{E}r(Y)$, its gradient is a bounded multiple of $\mathbb{E}[Gr(Y)]$. If $M \ge 1/2$ subtract 1 and factor $r-1 = (\sqrt{r}-1)(\sqrt{r}+1)$. Cauchy–Schwarz bounds $|\nabla M|^2/M$ by $C\mathbb{E}[|G|^2(\sqrt{r(Y)}-1)^2]$. If $M < 1/2$, bound it by $C\mathbb{E}[|G|^2 r(Y)]$ and use $\mathbb{E}(\sqrt{r(Y)}-1)^2 \ge (1-1/\sqrt{2})^2$ to absorb the constant term after $r \le 2(\sqrt{r}-1)^2+2$. Under $x\sim\gamma$, $\mathbb{E}[|G|^2 \mid Y] \le n+|Y|^2$. Integrating proves the claim. Apply it with $\lambda^2 = 1/(2w_*)$ to obtain (91). □

15.3 Rotation and the central interval

In the interior put $a=\cos\theta$, $b=\sin\theta$, $t=b^2$, $Z = aU+bV$ and $W = -bU+aV$. Let $\rho = \nabla \log p_Z$, $S = \rho + z$ and $H = DS$. For a vector field h, Dh is its Jacobian and $\operatorname{div}_p h = \operatorname{div} h + \rho\cdot h$. Set

$$m = \mathbb{E}[W \mid Z], \quad D = \mathbb{E}[WW^{\mathsf{T}} - I_n \mid Z], \quad Q = \operatorname{sym}\mathbb{E}[s_q(U)s_q(V)^{\mathsf{T}} \mid Z], \quad c = (a^2-b^2)/(ab).$$

Here $\operatorname{sym} A = (A+A^{\mathsf{T}})/2$, and $A:B$ is the Frobenius pairing. The rotational identities are

$$D = H + c\operatorname{sym} Dm + SS^{\mathsf{T}} + c\operatorname{sym}(Sm^{\mathsf{T}}) - Q/(ab), \tag{93}$$

$$4a^2b^2F_{X,tt} = I(p_Z) - \mathbb{E}[H:D] - \|\operatorname{div}_p m\|_2^2 + c\,\mathbb{E}[S\cdot m]. \tag{94}$$

For the first identity use the joint relative scores $A_0 = as_q(U)+bs_q(V)$, $B_0 = -bs_q(U)+as_q(V)$ in the z,w coordinates. Their derivatives satisfy $D_wB_0 - D_zA_0 = cD_zB_0$. Conditional integration by parts gives $D = \mathbb{E}[D_wB_0 + B_0B_0^{\mathsf{T}} \mid Z]$. Differentiating the conditional means S, m adds the respective conditional covariance terms. Symmetrize the result and use

$$B_0B_0^{\mathsf{T}} - A_0A_0^{\mathsf{T}} - c\operatorname{sym}(A_0B_0^{\mathsf{T}}) = -\operatorname{sym}(s_q(U)s_q(V)^{\mathsf{T}})/(ab).$$

For (94), the angle derivatives of the density are $-\operatorname{div}(p_Z m)$ and $\sum_{ij}\partial_{ij}(p_Z\mathbb{E}[W_iW_j \mid Z]) + \operatorname{div}(zp_Z)$. Since $\mathbb{E}Z\cdot W = 0$ and $\mathbb{E}\operatorname{tr}H = -I(p_Z)$, the chain rule for $t=\sin^2\theta$ gives the stated formula. These

calculations are justified by smooth cutoffs. Interior Gaussian smoothing of the joint law gives the derivative formulas and a Gaussian lower envelope for p_Z on compact angle intervals. Scores grow at most linearly, their derivatives are conditionally represented by bounded-order posterior moments, and the inputs have exponential moments. In particular $m, Dm, \operatorname{div}_p m$ have the indicated square integrability from conditional Jensen applied to joint scores, their derivatives and products.

Lemma 15.4. *If $p = p_q$ is a smoothing of a centrally symmetric isotropic log-concave vector and $h \in W^{1,2}(p)$ is an odd gradient vector field with $\operatorname{sym}\mathbb{E}[h(X_q)X_q^{\mathsf{T}}] = 0$, then*

$$\mathbb{E}|h|^2 \le (1/3 + O_n(q))\mathbb{E}\|Dh\|_{\mathrm{HS}}^2.$$

The same bound applies if h is odd and its entire moment matrix $\mathbb{E}[h(X_q)X_q^{\mathsf{T}}]$ vanishes.

Proof. Condition on the latent input X and expand $h(\sqrt{q}X + \sqrt{1-q}G)$ into Gaussian Hermites $e_\alpha(G)$, with vector coefficients $b_\alpha(X)$. Peel off one third of the conditional Gaussian Dirichlet energy. The remainder is bounded by the sum of $\mathbb{E}|b_\alpha|^2$ for $|\alpha| \le 2$, with factors $1 - |\alpha|/3$. The even degrees have mean zero by parity. Write $B_1 = (b_{\mathbf{e}_1}, \dots, b_{\mathbf{e}_n})$ for the degree-one matrix. It is symmetric when h is a gradient, by Gaussian integration by parts. The moment-matrix assumption yields

$$\mathbb{E}B_1 = -\sqrt{q/(1-q)}\,\operatorname{sym}\mathbb{E}[b_0(X)X^{\mathsf{T}}]$$

in the gradient case. In the full-orthogonality case the same identity holds without symmetrizing. Isotropy and Bessel give $\|\mathbb{E}B_1\|_{\mathrm{HS}}^2 \le q(1-q)^{-1}\mathbb{E}|b_0|^2$. Poincaré on X, followed by Bessel in G for Dh, bounds the sum of variances for $|\alpha| \le 2$ by $C_n q\,\mathbb{E}\|Dh\|_{\mathrm{HS}}^2$. Together with the Gaussian contribution $(1-q)\mathbb{E}\|Dh\|_{\mathrm{HS}}^2/3$, this proves the result. The argument extends from smooth Sobolev test fields by approximation (the mean-matrix estimate before the limit also includes the squared moment-matrix error). The gradient case uses cutoff and smoothing of the potential locally, or conditional Gaussian integration by parts for the weak Jacobian. □

Let $h = \nabla\psi$ be the $L^2(p_Z)$ projection of m onto gradients. It can be constructed by solving $\int \nabla\psi \cdot \nabla v\, p_Z = \int m \cdot \nabla v\, p_Z$ on the mean-zero weighted Sobolev space; Poincaré makes the gradient norm coercive. Then $L_{p_Z}\psi = \operatorname{div}_p m$ weakly and

$$\|\operatorname{div}_p m\|_2^2 = \|Dh\|_2^2 - \mathbb{E}[h^{\mathsf{T}} D\rho\, h].$$

For justification, local elliptic regularity gives $W^{2,2}_{\mathrm{loc}}$ for ψ. Cutoffs ψ_R on balls tending to infinity converge to ψ in $W^{1,2}(p_Z)$ and in the graph norm of L_{p_Z}. In the commutators, the score grows at most linearly and the cutoff has first derivatives of order R^{-1} and second derivatives of order R^{-2}. The terms with $\psi, \nabla\psi$ on the annuli therefore tend to zero. Mollifying each compact cutoff allows applying the weighted Bochner identity. The nonpositive $D\rho$ is bounded, so the identity on differences gives convergence of the Hessians and proves the formula for ψ.

Replacing m by h leaves the other pairings in (94) unchanged. Indeed S and $HS = \nabla(|S|^2/2)$ are gradients, and $\mathbb{E}H : Dm = -\mathbb{E}(L_{p_Z}S)\cdot m$, where $L_{p_Z}S$ is also a gradient. To see the latter, $S = \nabla\zeta$ and $L_{p_Z}S = \nabla(\Delta\zeta + \rho\cdot S) - (D\rho)S$, with $(D\rho)S = (H - I_n)S$ a gradient. The potentials here have at most polynomial growth and may be cut off in the weighted space. By reflection ψ is even and h is odd. Testing with gradients of quadratic polynomials gives $\operatorname{sym}\mathbb{E}[h(Z)Z^{\mathsf{T}}] = \operatorname{sym}\mathbb{E}[m(Z)Z^{\mathsf{T}}] = 0$. Also S is odd with its entire moment matrix zero.

On the span of S, h, Lemma 15.4 implies positivity of $\mathcal{E}(f,g) = \mathbb{E}[Df : Dg - f\cdot g]$, with $\mathcal{E}(f,f) \ge c\|Df\|_2^2$ for small q. The square in (94) is $\mathcal{E}(h,h) + \mathbb{E}[h^{\mathsf{T}}(2I_n - H)h]$. Complete the square in the nonlinear term:

$$-\mathbb{E}[S^{\mathsf{T}}HS + cS^{\mathsf{T}}Hh] \le \kappa\,\mathbb{E}|S + (c/2)h|^2 + (c^2/4)\mathbb{E}|h|^2.$$

This uses $-\kappa I_n \preceq H \preceq I_n$. For $|c| \le 2-\delta$, the excess of the Bochner square absorbs the last term. The quadratic form $\mathcal{E}(S,S) + c\mathcal{E}(S,h) + \mathcal{E}(h,h)$ controls $1-|c|/2$ times the sum of the two diagonal energies. The κ-term is absorbed for small q. By Jensen, $\|Q\|_2 \le I(p_q)$, so Lemma 15.3 and Poincaré bound its contribution by $O_{n,\delta}(q^{\eta_n} A(p_Z))$. We obtain $F_{X,tt} \le -c_n A(p_Z)$ uniformly on $1/5 \le t \le 4/5$.

15.4 Weak-score approximation at exterior weights

By swap symmetry take $a \ge b > 0$, $b^2 \le 1/5$. Put $v = q/2$, $d = \sqrt{2}b$, $\sigma_1 = 1/2$, $\sigma_2^2 = 1 - qa^2 - d^2 - \sigma_1^2$. With independent $Y \sim p_v$, X_1, G_1, G_2, represent the output by

$$R = dY + \sigma_1 G_1, \qquad Z = a\sqrt{q}X_1 + R + \sigma_2 G_2.$$

Both noise variances are bounded below. The weak score projection $J(z) = \mathbb{E}[s_q(V) \mid aU + bV = z]$ in this representation is $J(z) = \sqrt{2}\,\mathbb{E}[s_v(Y) \mid Z = z]$. Indeed, in the Fourier convention $+i\xi \cdot x$,

$$\widehat{s_u p_u}(\xi) = -i\sqrt{u}\, e^{-(1-u)|\xi|^2/2}[\nabla \chi_X(\sqrt{u}\,\xi) + \sqrt{u}\,\xi \chi_X(\sqrt{u}\,\xi)].$$

Inserting $u = q, v$ at gains b, d gives the same latent argument, and multiplication by the independent strong-arm and spare-noise factors gives the same total Gaussian factor. The assertion also applies to products of two signed scores with independent signal allocations.

Let e_α be the normalized multivariate Hermites. Put $A_\alpha(u) = \mathbb{E}[s_u(X_u) e_\alpha(X_u)]$. Coordinatewise integration by parts gives

$$[A_\alpha(u)]_i = \sqrt{\alpha_i + 1}\,\mathbb{E} e_{\alpha + \mathbf{e}_i}(X_u).$$

Thus $A_\alpha(q) = 2^{(|\alpha|+1)/2} A_\alpha(v)$, and degrees at most two and all even degrees vanish. Write $\|h\|_{1,Z}^2 = \mathbb{E}(|h(Z)|^2 + \|Dh(Z)\|_{\mathrm{HS}}^2)$.

Lemma 15.5. *For $b_* > 0$, $\varepsilon > 0$, there are N and a common small signal range such that, uniformly on $b_* \le b \le 1/\sqrt{5}$,*

$$\left\| J - \sum_{\substack{3 \le |\alpha| \le N \\ |\alpha| \text{ odd}}} b^{|\alpha|} A_\alpha(q) e_\alpha \right\|_{1,Z} \le \varepsilon \sqrt{I(p_v)}.$$

Jointly as $b, q \downarrow 0$, the same approximation restricted to $|\alpha| = 3$ has error $o(b^3)\sqrt{I(p_v)}$, uniformly over X.

Proof. The first channel, $h \mapsto \mathbb{E}[h(Y) \mid R]$, has conjugated Lebesgue L^2 kernel

$$K_d(x,y) = \sqrt{p_v(y)}\,\phi_{\sigma_1}(x - dy)/\sqrt{p_R(x)},$$

where ϕ_{σ_1} is isotropic normal of covariance $\sigma_1^2 I_n$. For $|y| > T$, move to $y_0 = y - \Delta y/|y|$. If the outward log slope on each ray is at most $-M$ beyond radius $T - \Delta > 0$, then $p_v(y_0) \ge e^{M\Delta} p_v(y)$. Integrate the denominator on the half unit ball around y_0 where $\nabla \log p_v(y_0) \cdot (u - y_0) \ge 0$. The Hessian lower cap $D^2 \log p_v \succeq -(1-v)^{-1} I_n$ gives $p_v(u) \ge e^{-C} p_v(y_0)$ there, with a fixed positive integration volume. Completing the square in the Gaussian kernel ratio yields

$$K_d(x,y) \le C_n e^{-M\Delta/2 + Cd^2(\Delta+1)^2} e^{-c|x-dy|^2}.$$

Schur's bound for the tail in y is therefore $C_n d^{-n/2} e^{-M\Delta/2 + Cd^2(\Delta+1)^2}$. On compact positive gains, take $\Delta = \eta T$ with small fixed $\eta > 0$. At fixed large T, the log slope on the sphere $T - \Delta$ converges uniformly to the Gaussian slope as $q \to 0$, by the local posterior bounds. The slope bound extends on each ray by log-concavity, so $M = cT$ is available for small q. This gives a Gaussian small tail in T. On compact $|y| \le T$, the kernel converges in Hilbert–Schmidt norm to the normal-input kernel. For domination use $p_R \ge c_n \phi_{\sigma_1}$, by averaging symmetric shifts and restricting $|Y| \le 2\sqrt{n}$. This gives a uniform squared Gaussian majorant in x for bounded y. Local convergence of densities follows from the natural coupling $Y \to G$ and the Gaussian kernel. Together these facts give operator-norm convergence to the normal channel on a compact positive gain range.

With $r = (d^2 + \sigma_1^2)^{1/2}$, the normal channel sends $e_\alpha(Y)$ to $(d/r)^{|\alpha|} e_\alpha(R/r)$, and $d/r < 1$ has a uniform margin. Truncate that representation by total degree. The finite basis vectors in the conjugated operator can use the actual densities in place of the normal densities, by convergence of the densities locally and uniform high moments.

The second channel from R to Z maps L^2 boundedly into $\|\cdot\|_{1,Z}$. The derivative of its conditional map is $-\operatorname{Cov}(h(R), G_2 \mid Z)/\sigma_2$. The Gaussian G_2 under this conditioning has covariance bounded above by I_n, by the Brascamp–Lieb variance inequality [14]: its conditional density is Gaussian times a log-concave factor. Rowwise conditional Cauchy–Schwarz proves the derivative estimate. More generally the same pointwise bound and Jensen give the norm bound in L^p for fixed $p \ge 2$. For each fixed α the second map sends $e_\alpha(R/r)$ to a quantity converging in $\|\cdot\|_{1,Z}$ to $r^{|\alpha|} e_\alpha(Z)$. The Gaussian kernels give local convergence of the conditional ratio and its derivative by the natural coupling, uniform moments and the lower bound on σ_2. The L^p estimates for $p > 2$ and tightness control the exterior z integrals. The limit also holds jointly as $d \to 0$. Composing the two maps gives the stated finite sums because $\sqrt{2}\, d^{|\alpha|} A_\alpha(v) = b^{|\alpha|} A_\alpha(q)$.

For the endpoint, take $T = d^{-1/2}$, $\Delta = T/3$ in the Schur estimate, with constant $M > 0$ from the slope at a fixed large radius. The tail norm is then smaller than any power of d. On $|y| \le T$ expand $\phi_{\sigma_1}(x - dy)$ through total degree three. Its fourth-order translated derivatives have bounded $L^2(p_R^{-1} dx)$ norms for $|dy| \le \sqrt{d}$. The remainder is $O(d^4)\|h\|_{L^2(p_v)}$ using the eighth moment, and restoring the tail to the coefficients of the polynomial terms has smaller cost by exponential moments. Applied to s_v, only degree three remains. Indeed all lower moments vanish, so $Y^\alpha/\sqrt{\alpha!}$ and $e_\alpha(Y)$ have the same score pairing at $|\alpha| = 3$. The resulting cubic is asymptotic in $L^2(p_R)$, up to $o(d^3)\sqrt{I(p_v)}$, to

$$(d/\sigma_1)^3 \sum_{|\alpha|=3} A_\alpha(v) e_\alpha(R/\sigma_1).$$

To replace p_R by ϕ_{σ_1} in the cubic ratio use compact convergence, $p_R \ge c_n \phi_{\sigma_1}$ and uniform high moments for the polynomial limit. Apply the second map and $r \to \sigma_1$ to get the second assertion. The first channel alone gives $\|\mathbb{E}[s_v(Y) \mid dY + sG]\|_2 \le C_{n,s} d^3 \sqrt{I(p_v)}$ for any other fixed $s > 0$ by the same proof. □

15.5 Closure at exterior weights

Keep $a \ge b$, $b^2 \le 1/5$. Now $m = -bS/a + J/a$, so the quadratic expression in (94), after discarding the nonpositive divergence square, is

$$-k\,\mathcal{E}(S,S) - (c/a)\mathcal{E}(S,J), \qquad k = b^2/a^2.$$

Here $\mathcal{E}(S,S) \ge c_n A(p_Z)$ by Lemma 15.4. The coefficient c/a is nonnegative in this region. We therefore only need a lower bound on the cross pairing. Write $K = A(p_Z)$. For each fixed finite

Hermite degree,

$$\mathbb{E}[S_i e_\alpha(Z)] = (a^{|\alpha|+1} + b^{|\alpha|+1})[A_\alpha(q)]_i + O_{n,\alpha}(K),$$
$$\mathcal{E}(S, \mathbf{e}_i e_\alpha) = (|\alpha| - 1)\mathbb{E}[S_i e_\alpha(Z)] - \mathbb{E}[S_i S \cdot \nabla e_\alpha(Z)].$$

For the first identity use Gaussian score integration by parts and the Hermite addition formula. Each additional term is a product of two positive-degree input Hermite moments, each $O_{n,\alpha}(\sqrt{K})$ by the score moments at v, (91), and the fixed signal ratio q/v. The Hermite addition formula is exact without any such term when $|\alpha| = 3$ by vanishing of the first three density moment differences. The second identity is weighted integration by parts and the Hermite equation. Applying Poincaré to $|S|^2$ gives $\|S\|_4^2 \le I(p_Z) + C_n\|DS\|_4^2 \le C_n q\sqrt{K}$ by Lemma 15.2 and $I(p_Z) \le C_n K$. Consequently $\|HS\|_2 \le C_n q\sqrt{K}$ and the polynomial error in the second identity is $O_{n,\alpha}(q\sqrt{K})$.

By Lemma 15.5, on fixed compact exterior weights the cross pairing has leading summands

$$(|\alpha| - 1)b^{|\alpha|}(a^{|\alpha|+1} + b^{|\alpha|+1})|A_\alpha(q)|^2 \ge 0.$$

For the finite sum the error is $O_{n,N}(qK)$, and the rank approximation costs at most $C_n\varepsilon K$ by (91) and $\|S\|_{1,Z} \le C_n\sqrt{K}$. Choose ε and then the signal range to make the resulting error negligible relative to b^2K. For $b, q \downarrow 0$ use the cubic approximation instead, giving $\mathcal{E}(S, J) \ge -o(b^3K)$ and $\|J\|_{1,Z} \le C_n b^3\sqrt{K}$. Since $c/a = O(1/b)$, the cross term still costs $o(b^2K)$.

The nonlinear terms in (93) have $SS^\mathsf{T} + c\,\mathrm{sym}(Sm^\mathsf{T}) = kSS^\mathsf{T} + (c/a)\,\mathrm{sym}(SJ^\mathsf{T})$. In (94) their contributions are bounded above by $k\kappa I(p_Z) + Cb^{-1}\|HS\|_2\|J\|_2$. This is $O_n(qb^2K)$ near $b = 0$ and $O_n(qK)$ away from zero. The Q-term on compact exterior weights is $O_n(q^{\eta_n}K)$ by Lemma 15.3. Near zero, allocate the two signed scores at $v_h = 5q/4$ and $v = q/2$. The squared gains $a^2/(5/4)$ and $2b^2$ sum to less than one with a fixed margin. Use a fixed share of the remaining Gaussian variance for the weak channel. Condition first on its output and the strong independent input. The endpoint channel bound from Lemma 15.5, independence and conditional contraction then give

$$\|Q\|_2 \le C_n b^3\sqrt{I(p_v)I(p_{v_h})} \le C_n q b^3\sqrt{K}.$$

The factor $1/(ab)$ costs only $O_n(1/b)$. Thus all errors are absorbed by the term $-k\mathcal{E}(S, S)$, first in a fixed small endpoint window and then on the remaining compact range with a common sufficiently small q_n. Dividing by $4a^2b^2$ proves the theorem throughout the open weight interval.

At an endpoint use the parametrization in $b = \sqrt{t}$, which makes the entropy even in b by central symmetry. The entropy is C^4 near $b = 0$: the output has a fixed nondegenerate Gaussian noise and the gains $\sqrt{q}\sqrt{1 - b^2}, \sqrt{q}\,b$ are smooth there. Differentiating the Gaussian kernel represents each density derivative divided by the density as a conditional polynomial in the latent variables and Gaussian innovation, with bounded coefficients on a small parameter interval. Jensen and the uniform moments give integrable tail bounds for the products of such ratios in the first four entropy derivatives. The term with $\log p$ grows at most quadratically by the Gaussian lower envelope. This justifies the derivatives and their continuity by local convergence and tail truncation. The fourth derivative in b at zero, divided by 12, is the continuous one-sided weight curvature. Also $A(p_Z)$ is continuous at the endpoint by local kernel convergence, the Hessian cap and tail tightness. The inequality therefore holds there as well. If $A(p_Z) = 0$, then Z is standard normal. On every line, the odd cumulants of X vanish by symmetry and each even higher cumulant of Z is the corresponding cumulant of X times $q^{j/2}(a^j + b^j) > 0$. Exponential moments identify each one-dimensional projection as standard normal on unit directions, and hence X is standard normal.

15.6 Consequences

The largest universal signals are nonincreasing in dimension, by adjoining independent Gaussian coordinates. For product laws with independent symmetric variance-one log-concave coordinates, entropy additivity gives a signal range independent of dimension directly from Section 14. An invertible affine change of variables preserves the curvature if the Gaussian used for smoothing has the same covariance as the latent law. Whitening therefore gives the corresponding statement for centrally symmetric log-concave laws with any positive definite covariance.

A Uniform endpoint restoration for log-concave inputs

Fix a symmetric variance-one log-concave input and $0 < q \le 1/20$. Let $U = X_q$, with density p, score $\rho = (\log p)'$, and put

$$g = -\rho', \quad s = \rho + x, \quad A = \mathbb{E}(g(U) - 1)^2, \quad \delta = \mathbb{E}s(U)^2, \quad I = \mathbb{E}\rho(U)^2, \quad K = \mathbb{E}g(U)^2.$$

The letters g, δ, K in this section are score quantities, distinct from the relative-density notation in Section 2.

Gaussian convolution preserves log-concavity, and the posterior-variance identity for the Gaussian kernel gives

$$0 \le g \le M := \frac{1}{1-q}. \tag{95}$$

Indeed $(\log p)'' = -1/(1-q) + q\,\mathrm{Var}(X \mid U)/(1-q)^2$, while log-concavity supplies the opposite bound. The score is odd and $|\rho(x)| \le M|x|$.

The exact symmetric endpoint identity from [2, Theorem 3.1] is

$$F_{tt}(q, 0) = -\frac{A}{2} - \frac{\kappa_4(U)}{12}B, \qquad B = \mathbb{E}g^2 - \frac{1}{3}\mathbb{E}\rho^4. \tag{96}$$

We first justify its use without original input smoothness. Put $t = u^2$ in (1) and allow $|u| \le 1/4$. The Gaussian-kernel argument of Lemma 3.1 applies to the smooth gain functions $\sqrt{q}\sqrt{1-u^2}$, $\sqrt{q}\,u$, with fixed positive noise variance. It proves $H(u) = F_X(q, u^2)$ is C^4 and even. Hence

$$\lim_{t \downarrow 0} F_{tt}(q, t) = H^{(4)}(0)/12.$$

The fourth-order density expansion is

$$\begin{aligned} p_u &= p + c_2 u^2 + c_4 u^4 + o(u^4), \\ c_2 &= (p + xp' + p'')/2, \\ 2c_4 &= 3p/4 + 5xp'/4 + x^2 p''/4 + 3p''/2 + xp'''/2 + \mathbb{E}U^4\, p''''/12. \end{aligned}$$

Differentiating entropy gives $F_{tt}(q, 0) = -\int 2c_4 \log p - \int c_2^2/p$. The Gaussian fourth-moment portion is the normalized Ornstein–Uhlenbeck curvature $-A/2$; the remaining term is $-\kappa_4(U) \int p'''' \log p/12$. Two integrations by parts and $\mathbb{E}[\rho^2\rho'] = -\mathbb{E}\rho^4/3$ give (96). This reproduces the cited identity and its OU decomposition. Smooth cutoffs justify these integrations: the conditional-polynomial derivative bounds control spatial derivatives, g is bounded, ρ is at most linear, and all input moments are finite.

We use the standard one-dimensional Poincaré bound $C_P(X) \le 12\,\mathrm{Var}(X)$ for log-concave measures [3]. Tensorization applied to $f(\sqrt{q}X + \sqrt{1-q}G)$ gives

$$C_P(U) \le qC_P(X) + (1-q) \le 1 + 11q =: C \le 31/20 < 8/5.$$

Integration by parts gives

$$I = \mathbb{E}g = 1 + \delta, \qquad \mathbb{E}[Us(U)] = 0, \qquad K = 1 + A + 2\delta.$$

Poincaré applied to the odd function s yields $\delta \le CA$. Also $g^2 \le Mg$, and $M \le 2, I \ge 1$, so

$$A = K - 2I + 1 \le (M-2)I + 1 \le M - 1 \le 1/19 < 1/16,$$

and

$$K \le 1 + (1 + 2C)A \le 116/95 < 5/4. \tag{97}$$

Write κ_j for cumulants. The identity $\mathbb{E}[U^3\rho(U)] = -3$ implies

$$\kappa_4(U) = \mathbb{E}[(U^3 - 3U)s(U)].$$

Cauchy–Schwarz gives

$$|\kappa_4(U)| \le \sqrt{\mathbb{E}[(U^3 - 3U)^2]CA}.$$

The exact moment expression is

$$\mathbb{E}[(U^3 - 3U)^2] = 6 + 9\kappa_4(U) + \kappa_6(U).$$

Cumulant addition gives $\kappa_4(U) = q^2\kappa_4(X)$ and $\kappa_6(U) = q^3\kappa_6(X)$. The moment bound $\|X\|_p \le p/2$ for $p \ge 2$ [9] gives $\mathbb{E}X^4 \le 16$, $\mathbb{E}X^6 \le 729$, hence $\kappa_4(X) \le 13$ and $\kappa_6(X) \le 759$. Therefore

$$\mathbb{E}[(U^3 - 3U)^2] \le 6 + 117q^2 + 759q^3 < 7, \quad |\kappa_4(U)| < \frac{27}{8}\sqrt{A}.$$

For clarity we reproduce the centered correction estimate from the endpoint framework [2]. Let $D_0 = g - \rho^2$, with $\mathbb{E}D_0 = 0$. Integration by parts gives $\mathbb{E}[g\rho^2] = \mathbb{E}\rho^4/3$, so

$$B = \mathbb{E}[(g-1)D_0], \qquad \mathbb{E}D_0^2 = 2K - B, \qquad B^2 \le A(2K - B).$$

Consequently, for $A > 0$,

$$\frac{|B|}{\sqrt{A}} \le \frac{\sqrt{A} + \sqrt{A + 8K}}{2} < \frac{7}{4},$$

by $A < 1/16$ and (97). Combining the two bounds gives $|\kappa_4(U)B| < (189/32)A$. Equation (96) now yields

$$F_{tt}(q, 0) \le -\frac{6 - 189/32}{12}A = -A/128.$$

If $A = 0$, smooth positivity implies $(\log p)'' = -1$ everywhere, and centering and normalization give $p = \varphi$. Dividing characteristic functions by the nonzero Gaussian factor shows X is Gaussian. Otherwise the bound is strict. The symmetry in the mixing weight supplies the other endpoint, completing Theorem 1.8.

B Moment-matched exponential tails

The Gaussian-relative and weak/moment neighborhoods are distinct. A tangent-tail potential and compactly supported moment corrections give the following explicit separating family.

Theorem B.1. *For every integer $k \geq 1$ there are symmetric variance-one log-concave densities f_R, with $R \to \infty$, whose Gaussian moments agree exactly through order $2k$, and whose first unmatched cumulant satisfies*

$$\kappa_{2k+2}(f_R) \sim 2\varphi(R)R^{2k-1} > 0.$$

Their laws converge to the Gaussian in total variation and in every fixed absolute moment. Nevertheless, for each member and every $q \in (0, 1]$, the density of its Gaussian smoothing has an unbounded ratio to φ and infinite chi-square divergence from φ.

Proof. Let $V_R(x) = x^2/2$ for $|x| \leq R$ and $V_R(x) = R|x| - R^2/2$ otherwise, and let $b_R = (2\pi)^{-1/2}e^{-V_R}$. This potential is even, convex and C^1. For fixed integer $j \geq 0$, put $B_{j,R} = \int x^{2j}(b_R - \varphi)$ and $A_R = B_{0,R}$. Substitution $x = R + y/R$ in both tails gives

$$\begin{aligned} B_{j,R} &= 2\varphi(R)R^{2j-1} \int_0^\infty e^{-y} \left(1 + \frac{y}{R^2}\right)^{2j} \left(1 - e^{-y^2/(2R^2)}\right) dy \\ &\sim 2\varphi(R)R^{2j-3}. \end{aligned}$$

Dominated convergence after multiplying the integral by R^2 uses the majorant $e^{-y}(1 + y)^{2j}y^2/2$ for $R \geq 1$; the limiting integral is one. Thus $A_R \sim 2\varphi(R)R^{-3}$. If $m_j = (2j - 1)!!$, the moment deviations of the normalized base law are

$$d_{j,R} = \frac{B_{j,R} - m_j A_R}{1 + A_R} \sim 2\varphi(R)R^{2j-3} \quad (j \geq 1).$$

Choose an even smooth cutoff χ_a, equal to one on $[-a, a]$ and zero outside $[-a - 1, a + 1]$, and set $\psi_j(x) = \chi_a(x)x^{2j}$ for $j = 1, \ldots, k$. For sufficiently large fixed a, the matrix

$$C_{ij} = \mathrm{Cov}_\gamma(G^{2i}, \psi_j(G)), \qquad 1 \leq i, j \leq k,$$

is invertible. Indeed it converges as $a \to \infty$ to the positive-definite covariance matrix of $G^2, \ldots, G^{2k}$: a nonconstant polynomial cannot be almost surely constant under a positive Gaussian density. Freeze such controls.

For $R > a + 1$ tilt the base density by $e^{-\sum \theta_j \psi_j}$ and normalize, and let $M_R(\theta)$ be its vector of moment deviations of orders $2, \ldots, 2k$. Then $M_R(0) = d_R$ and

$$(DM_R(\theta))_{ij} = -\mathrm{Cov}_{R,\theta}(X^{2i}, \psi_j(X)).$$

Boundedness of the controls and the base moment estimates imply uniform convergence of these derivatives on a fixed small parameter ball to those for the corresponding tilted Gaussian, and give uniform bounds on second derivatives. In particular $DM_R(0) \to -C$ has uniformly bounded inverse. Writing $M_R(\theta) = d_R + J_R\theta + E_R(\theta)$, the map

$$\theta \longmapsto -J_R^{-1}(d_R + E_R(\theta))$$

is a contraction of a ball of radius $C_0 \|d_R\|$ into itself for large R. Its fixed point θ_R gives exact moment matching and

$$\|\theta_R\| = O(\varphi(R)R^{2k-3}).$$

This is the usual finite-dimensional implicit-function correction, with a uniform inverse and quadratic remainder.

The corrected potential is $W_R = V_R + \sum \theta_{j,R}\psi_j$. On the fixed control support, $V_R'' = 1$ and $W_R'' \geq 1/2$ for large R; outside it the potential equals the convex V_R. The smooth controls have

vanishing value and first derivative at their support boundary, so W_R is globally convex. The density is therefore log-concave, even, and already has variance one by exact matching. Its tails retain rate R.

For every fixed moment order, the correction changes its value by $O(\|\theta_R\|)$, by the bounded-control covariance formula along the parameter segment. At order $2k+2$, this is smaller by R^{-2} than the base deviation $2\varphi(R)R^{2k-1}$. The stated first-unmatched-moment asymptotic follows; because all lower moments match, that deviation is also the cumulant. The tilt converges uniformly to one. The same tail-integral argument works for every fixed nonnegative absolute power and proves moment and total-variation convergence to the Gaussian.

Finally the corrected density has an exact positive exponential tail $c_R e^{-Rx}$ for large positive x. In its Gaussian convolution at $q < 1$, restrict the Gaussian summand to absolute value at most one. For all sufficiently large y this yields

$$p_{R,q}(y) \geq C_{R,q} e^{-Ry/\sqrt{q}}.$$

For $q = 1$ the tail identity gives the same conclusion directly. Therefore $p_{R,q}/\varphi$ is unbounded, and $\int p_{R,q}^2/\varphi$ diverges by comparison with a positive constant times $\exp(y^2/2 - 2Ry/\sqrt{q})$ on a tail. □

Statements and Declarations

Funding. This research was self-funded. No external research funding was received.

Competing interests. The author owns Luminary Oddities. The author has declared no additional relevant competing interests.

Data availability. All mathematical arguments are contained in this article.

AI assistance. ChatGPT (OpenAI, Sol 6.1) assisted with literature research, mathematical exploration, proof development and checking, and manuscript preparation. Scott Schubr directed the research and manuscript development and retains responsibility for the content, accuracy, originality and attribution of the work. The sections added in the sixth through eighth editions were developed with OpenAI Codex (derivation, source checks and independent review) and edited and cross-referenced with Anthropic Claude.

References

[1] E. Abbe and L. Zheng, *A coordinate system for Gaussian networks*, arXiv:1002.2271, 2010. `https://arxiv.org/abs/1002.2271`.

[2] F. Abu-Ajamieh, *Noise-Induced Restoration of Entropy Concavity Near Gaussian Equilibrium*, arXiv:2610.22386v1, 2026. `https://arxiv.org/abs/2610.22386v1`.

[3] P. Cattiaux, A. Guillin and P.-A. Zitt, *Poincaré inequalities and hitting times*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **49** (2013), no. 1, 95–118, doi:10.1214/11-AIHP447; see equation (4.1). `https://www.numdam.org/item/10.1214/11-AIHP447.pdf`.

[4] A. Eskenazis and L. Gavalakis, *On the entropy and information of Gaussian mixtures*, Mathematika, 2024, doi:10.1112/mtk.12246.

[5] Anonymous, *A sharp smoothing inequality in Gauss space and the exact threshold for local entropy concavity*, unrefereed AI-generated candidate, revised November 11, 2026, doi:10.5281/zenodo.24144687. `https://github.com/ipitchford/sharp-smoothing-entropy-threshold`.

[6] D. Guo, Y. Wu, S. Shamai and S. Verdú, *Estimation in Gaussian noise: Properties of the minimum mean-square error*, arXiv:1004.3332, 2010. `https://arxiv.org/abs/1004.3332`.

[7] C. Luo, *Near-Gaussian counterexamples to the Ball–Nayar–Tkocz entropy concavity conjecture*, arXiv:2609.11418v2, 2026. `https://arxiv.org/abs/2609.11418v2`.

[8] M. Madiman, P. Nayar and T. Tkocz, *Two remarks on generalized entropy power inequalities*, in Geometric Aspects of Functional Analysis: Israel Seminar (GAFA) 2017–2019, Volume II, B. Klartag and E. Milman (eds.), Lecture Notes in Mathematics 2266, Springer, 2020, doi:10.1007/978-3-030-46762-3_7; arXiv:1904.02314v1 (April 4, 2019). `https://arxiv.org/abs/1904.02314v1`.

[9] D. Murawski, *Comparing moments of real log-concave random variables*, Bernoulli **32** (2026), no. 3, 2403–2426, `https://doi.org/10.3150/25-BEJ1958`; arXiv:2211.05210. `https://arxiv.org/abs/2211.05210`.

[10] A. Saumard and J. A. Wellner, *Log-concavity and strong log-concavity: a review*, Statistical Surveys **8** (2014), 45–114. `https://arxiv.org/abs/1404.5886`.

[11] E. Nelson, *The free Markoff field*, J. Funct. Anal. **12** (1973), 211–227.

[12] NIST Digital Library of Mathematical Functions, Chapter 18, Equations 18.12.13 and 18.15.14–16. Generating function and fixed-argument asymptotic.

[13] C. L. Frenzen and R. Wong, *Uniform asymptotic expansions of Laguerre polynomials*, SIAM J. Math. Anal. **19** (1988), 1232–1248. DOI.

[14] H. J. Brascamp and E. H. Lieb, *On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems*, J. Funct. Anal. **22** (1976), 366–389. Journal paper.

[15] K. Ball and V. H. Nguyen, *Entropy jumps for isotropic log-concave random vectors and spectral gap*, Studia Math. **213** (2012), 81–96. Journal paper.

[16] S. G. Bobkov, *Isoperimetric and analytic inequalities for log-concave probability measures*, Ann. Probab. **27** (1999), 1903–1921. Journal paper.