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. As $q \to 1$ the endpoint interval carrying a uniform concavity neighborhood has sharp leading width $\sqrt{1-q}$, with moment-matched strongly log-concave witnesses of positive curvature beyond it. For each fixed centrally symmetric isotropic input, strong smoothing restores strict concavity on the whole simplex with the correct face factors, while for three or more copies moment-matched counterexamples occur at a fixed interior weight at every positive signal; the minimum trim for local concavity has sharp order q^2. The exact region of Gaussian input loads for the binary conditional-product estimate is determined in every fixed dimension, and after strong smoothing all isotropic log-concave laws have strictly concave mixing entropy on every weight region with minimum weight at least $q^2[\log(1/q)]^B$, where the power two is sharp. Aligned oscillator modes give positive curvature at prescribed unequal weights, improving the high-signal endpoint width bound to $\sqrt{1-q} - \frac{7}{12}(1-q) + O((1-q)^{3/2})$. For a log-concave law matching Gaussian moments through order $n-1$, the normalized nth cumulant has sharp

exponential rate $2^{-n/2}$, with explicit dependence on the density-ratio radius and both signs realized by smooth strongly log-concave densities. Every fixed isotropic law with first moment mismatch at order $n \geq 3$ and finite moment of order $2n+\varepsilon$ eventually has strictly concave mixing entropy on the whole simplex. For centered variance-one log-concave laws with moments fixed through order $n-1$, the attainable nth moments form a closed interval whose unique extremizers are piecewise log-affine; at zero skew the extremal fifth moment is attained by a two-piece exponential.

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 fourteenth edition of this article. The fourteenth edition adds Sections 23–25, which prove eventual whole-simplex concavity for fixed asymmetric inputs and solve the log-concave signed moment problem, including the extremal fifth moment at zero skew. The thirteenth edition added Section 22, which determines the sharp order of the first unmatched Gaussian cumulant of a log-concave law, over the whole class and in density-ratio neighborhoods. The twelfth edition added Section 21, which constructs positive curvature at prescribed unequal weights by aligned oscillator modes and sharpens the high-signal endpoint width to second order. The eleventh edition added Sections 19 and 20, which determine the exact binary Gaussian conditional-product region in every fixed dimension and prove whole-class log-concave concavity on expanding trimmed simplices. The tenth edition added Sections 17 and 18, which prove fixed-input eventual concavity on the whole simplex with moment-matched counterexamples at a fixed interior weight, and the sharp trimming scale q^2 for multicopy concavity. The ninth edition added Section 16, which determines the sharp leading width $\sqrt{1-q}$ of the uniform endpoint concavity interval as the retained signal tends to one. The eighth edition added 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

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

Given independent copies U_q, V_q of X_q, set

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

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 $m \geq 3$ independent copies $X_1, \ldots, X_m$ and an independent standard Gaussian G, define on the probability simplex

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

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

Theorem 1.1 (Moment-matched three-copy obstructions)**.** *Fix $q \in (0,1)$ and an integer $K \geq 2$. 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 $\|g_j\|_{C_b^r} \to 0$ for every fixed integer $r \geq 0$, 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 $\|g\|_{C_b^r}$ 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 $\sum_i v_i = 0$,

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

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

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

while $F_{X_j}(q, \cdot)$ fails concavity for every j. Here C_b^m 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:

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

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 $N \ge 4$, $B < \infty$ and $\varepsilon > 0$. Let $\mathcal{C}$ be a family of symmetric variance-one laws satisfying*

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

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

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

There is $q_{B,N,\varepsilon} > 0$ such that every $X \in \mathcal{C}$ has strict all-weight entropy concavity for every $0 < q \le q_{B,N,\varepsilon}$, 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: $\sup_{X\in\mathcal{C}} \mathbb{E}|X|^j \le B_j < \infty$ for every integer $j \ge 1$. For every fixed even $N \ge 4$, there are $K_N < \infty$ and $q_N > 0$ such that, for every $X \in \mathcal{C}$ and $0 < q \le q_N$,*

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

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 $0 < q \le q_{P,A}$, implies

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

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 $n = 2k \ge 4$, its moments below order n equal the Gaussian moments, $\kappa_n = \mathbb{E}X^n - \mathbb{E}G^n \ne 0$, and $\mathbb{E}|X|^{2n+2} < \infty$. There is $q_X > 0$ such that $\partial_t^2 F_X(q,t) < 0$ for every $0 < q \le q_X$ 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 $n = 2k \ge 4$ 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 $S_k \ge S_{k-1}/2$ and $S_k \ge S_{k-2}/4$; also $S_{k-2} \ge S_k$. Thus

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

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

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

The inverse is restricted to positive Hermite degrees. If $g \ne 0$, 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 $S_k(t) \geq 2^{1-k}$,

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

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

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

It follows that $\|NM(q,z)\|_2 \leq K\delta \|L\|_2 / 2$. 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

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

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 $0 < q \leq Q$.

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

$$p_H(z) = \mathbb{E}\varphi_{\sigma(H)}(z - h_1X - h_2Y).$$

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 $u = (z - h_1X - h_2Y)/\sigma$. 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

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

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

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

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 $k \geq 2$,

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

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

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

for all $0 < q \leq q_X$ and all weights. This proves Theorem 1.7. For the all-moment assertion, Gaussian moments are determinate: $(2j-1)!! \leq (2j)^j$ 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 $|H| \leq 1/2$, Chebyshev's inequality gives $\Pr(|X| \leq 2) \geq 3/4$. 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

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

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

$$L_n = a_n(X)(a^{n/2} + b^{n/2}),$$

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

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 $(S_k^2)'' \ge k(k-1)2^{5-2k} > 0$. Let

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

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

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

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

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

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

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

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

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

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

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

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

For each fixed even $n \ge 4$ and $A > 0$, choose one even $N \ge \max\{4, n, n + 2A - 1\}$ and take $q \le 1$ 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 $0 < q \le q_{\mathcal{U}}$. 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 $\|X\|_p \le p/2$ for $p \ge 2$ [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

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

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 $n = 2\ell \ge 6$ 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 $\kappa_n(U) = q^\ell \kappa_n(X) \ne 0$. 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

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

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

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

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

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

$$B_n(p) = \mathbb{E}_\varphi[\mathrm{He}_n\, r] + Q_n(r) + \text{cubic remainder},$$
$$Q_n(h) = \int (\varphi h)^{(n)} h - \tfrac{1}{2}\mathbb{E}_\varphi[\mathrm{He}_n\, h^2]. \tag{39}$$

Indeed $\int \varphi^{(n)} \log\varphi$ and $\int (\varphi r)^{(n)} \log\varphi$ vanish identically because $\log\varphi$ is quadratic and $n \geq 3$. The linear and quadratic terms left by $\log(1 + r)$ are exactly those displayed. For the later family $r = \alpha H + \beta\alpha^2 W$ 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

$$Q_n(T_{\sqrt{q}} e_d) = -\tfrac{1}{2} q^d c_{n,d} < 0. \tag{40}$$

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

7.3 Compact moment corrections and continuity

Let $I = \{0, 2, \ldots, n\}$. 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 $\mathbb{E}_\varphi \operatorname{He}_i w = 0$ for $i < n$ and $\mathbb{E}_\varphi \operatorname{He}_n w = 1$. 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

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

Here $\mathbb{E}_\varphi \operatorname{He}_n w = 1$ 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

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

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

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

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 $m \ge 3$. 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 $\|w - u_m\|_2 \le \rho$, 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 $0 < q \le q_m$ 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 $0 \le s \le \rho$ whenever $0 < \|g\|_\infty \le \delta$. 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 $m \ge 3$ 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 $\min_i w_i \ge \eta$, for every non-Gaussian admissible input and all $0 < q \le \eta^{3/2}/(256\sqrt{m})$.*

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

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

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 $1 \le j \le m$. 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 $j \ge 2$ when $q < q_m$. At equality it is strict for $2 \le j < m$. 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 $n = 2k \ge 4$ 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 $g \ne 0$. On fixed positive weight regions one has

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

whenever $a \le q\sqrt{(\min_i w_i)(\max_i w_i)}$. 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 $w = u_m + sv$, $s \ge 0$, a largest weight has $v_i = \max_j v_j \ge [m(m-1)]^{-1/2}$. This gives the stronger $E \ge c_m \|\mathcal{N}L\|_2^2$ there.

Put $\delta = \|g\|_\infty$. For real weights R and its first two spatial derivatives have uniform sup norm $O(\delta)$ for $q \le Q < 1$. Indeed the conditional Gaussian input vector has covariance $I - aa^\top$, $a_i = \sqrt{q w_i}$, 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 $\|R\|_\infty$,

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

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

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

using $\|R_{xx}\|_2 \le \|\mathcal{N}R\|_2$ 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

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

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

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

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 $b > \max_i q w_i$ close to a with the same estimate. The ratio $q w_i / b$ 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 $j \ge 2$ 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 $a \ge q/m$. The function $\tau_a(q)$ decreases with a, while λ_a increases, so $\tau_a(q)^m \le \lambda_a$, 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 $\|(\mathcal{N} + 1) T_{\sqrt{a}} g\| \le C D_*$ by the nonnegative linear multipliers and absence of low degrees. By conditional product contraction, commutation of $\mathcal{N}$ with total number and (51),

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

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 $\sum_{j \ne i} \mu_{\min(i,j),\max(i,j)} + 2\mu_{ii} = 1$. 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 $\|R\|_\infty \le e^{-\sqrt{n}} \operatorname{poly}(n)$, so this error is negligible compared with $n e^{-2m\sqrt{n}} \|B_{m,n}(u_m)\|_2^2$. 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 $q_3 = 0.65767078\ldots$ and $q_4 = 0.71210929\ldots$.

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

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

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 $i \in H$,

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

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,

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

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 $b \ge a$,

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

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 $\prod_i \tau_a(b_i) \le \lambda_a$, because the same inequality holds at u for $q \le q_4$. 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 $j \log \tau_{a_0}(ja_0)$ from $j = 1$ to m Sections 9.1–9.5. By (59), the bounds just proved for the full and proper subsets give

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

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

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 $b \ge a$. 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 $a_{\mathrm{ref}} \le a$. 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

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

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 $C\delta^2 X$. 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 $m \geq 5$, $q = q_m$, $a = q/m$, 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 $S_n(b) \leq C \max\{\nu_a(b), |\tau_a(b)|\}^n$. The maximum is $\nu_a(b)$ for $0 < b \leq a$ and $\tau_a(b)$ for $b \geq a$. 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

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

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

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

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 $\sqrt{a_- a_i} \leq b_\times$. Thus, with polynomial factors allowed in n, its magnitude is bounded by $\nu_a(b_\times)^n$. In particular

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

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

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

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,*

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

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

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

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 $d \ge 3$. 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, $w_i \ge 0, \sum_i w_i = 1$, and $0 < q < 1$. The barycentre is $u = (1/m, \dots, 1/m)$. Define q_m by $\tau_{q_m/m}(q_m)^m = \lambda_{q_m/m}$, with $q_2 = 4/7$.

Theorem 11.2. *Fix $d \ge 1$. For $m = 2$, one bounded Gaussian-relative neighborhood has strictly negative weight curvature for non-Gaussian admissible inputs at all mixing weights and all $0 < q \le 4/7$, 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 $\|e^{-cx^2/2} H_k\|_{L^2(dx)}^2 \leq C(k+1)^{-1/2}$. It follows that on a d-dimensional oscillator shell

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

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*

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

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

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

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

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

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,

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

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

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

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

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

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 $\|\mathcal{N}^{-1}L''\|_2 \le CE^{1/2}$ 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 $\max(t, 1-t) \geq 1/2$. 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

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

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 $\|\nabla^2 R\|_{L^2(\mathrm{HS})} \leq \|\mathcal{N}R\|_2$ 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

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

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

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

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 $\|\mathcal{N}\psi(L)\| \leq C\delta D_{a_{\mathrm{ref}}}$. 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 $w_{(2)}w_{(3)} \geq 1/9$, 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 $\prod_i \tau_a(b_i) \leq \lambda_a$ near u. Indeed $a \geq (q_m/4)(1 - 4\|w-u\|^2 + O(\|w-u\|^4))$ 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: $\|\mathcal{N}M\| + \|\mathcal{N}^{-1}M''\| \leq C\delta X_1$, 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

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

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 $a_{\mathrm{ref}} \leq a$. Part of the margin pays the number factors from weight derivatives. Conditional contraction and the strict scalar estimates thus give a product bound $C\delta D^2_{a_{\mathrm{ref}}}$ 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 $m \geq 5$ 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

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

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 $q w_i / c_i$. 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

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

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 $C Y^{2j-2} \|(\mathcal{N}+1)\mathcal{T}_a g\|_2^2$. 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

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

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 $M = \sum_{|H|\ge 2} B_H(g)$. 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 $E \ge cX_*^2$ and $\|\mathcal{N}^{-1}L''\| \le CX_*$, 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 $q = 1/2$, $q^2 \le a$ at every weight. For $m \ge 3$ at $q = q_0$, the full subset has $q_0^m = q_0/m \le a$. At u every proper subset has strict $(q_0 j/m)^j < q_0/m$, $2 \le j < m$. 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 $\|(\mathcal{N}+1)\mathcal{T}_a g\|_2 \le C X_*$ by the nonnegative linear multipliers,

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

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 $\|\mathcal{T}_{b'} g\|_2^2 \le \|g\|_\infty \|\mathcal{T}_a g\|_2$. 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 $1/2 \le q \le Q$. When $\min(t, 1-t) \le \eta$, 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 $4q \min(t, 1-t) \le 1$ gives $c_1 c_2 \le a$. 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 $\max(t, 1-t) \ge q + \Delta$, 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 $\|R\|_\infty \le (1 + \|g\|_\infty)^m - 1$. 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

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

To see this, Gaussian integration by parts gives $\||\nabla R|^2\|_2 \le 3 \|R\|_\infty \|\mathcal{N} R\|_2$. Integrate the Gaussian divergence of $R|\nabla R|^2 \nabla R$ and use $\|\nabla^2 R\|_{L^2(\mathrm{HS})} \le \|\mathcal{N} R\|_2$. 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 $-E \le -c X_*^2$. 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 $(c_1c_2)^k \le a^k$, $c_1 \le 1$, 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 $N^2 e^{-cN} c_1^{n_1} c_2^{n_2} \le Ca^k$.

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 $\max(t, 1-t) \geq 1/2$. Consequently $\langle L'', L\rangle \geq cX^2$ and $\|\mathcal{N}^{-1}L''\|_2 \leq CX$. For a smooth bounded Gaussian function f, integrating the Gaussian divergence of $f|\nabla f|^2\nabla f$ proves

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

using $\|\nabla^2 f\|_{L^2(\mathrm{HS})} \leq \|\mathcal{N}f\|_2$ and Cauchy–Schwarz. The Gaussian number chain rule thus gives $\|\mathcal{N}\psi(R_0)\|_2 \leq C\|R_0\|_\infty\|\mathcal{N}R_0\|_2$. 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 $0 < q \leq 1$ and let g be an admissible perturbation as defined above, with $\|g\|_\infty \leq 1/10$. If $q = 1$, assume $\mathcal{N}g \in L^2(\gamma_d)$. Then F'' extends continuously from the interior to both endpoints, and*

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

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 $J \geq 4$ occurs for $g \otimes g$. For $0 < t \leq 1/2$ and $\mu \leq 1$, the squared numerators on $J \geq 4$ are bounded by $CJ^4\mathbf{1}_{J\geq 4} \leq C(J)_4$, whose binomial expectation is at most $C\mu^4$. For $\mu \geq 1$, 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

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

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 $\min(t, 1-t) \leq 1/M$, since $M \geq 4$ 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

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

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,*

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

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 $\Pr(|X_q| > x_0) \le C e^{-c/q}$ 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 $|s_q'(y)| \geq q$. The above score bounds give $p_q(x) \leq e^{-c'/q} p_q(y)$ along either shift. The shift is injective on each tail and their images are disjoint. Hence the entire set $s_q' \geq Kq$ has mass at most $q^{-2}e^{-c'/q}A(p_q)$. On its complement, $|s_q'| \leq (K+2)q$. 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,*

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

is a convex function of $\Theta(u)$ at all sufficiently large u, uniformly over the class, where $\Theta'(u) = \exp(-C_0 e^{-2u})$. 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 $A^2 \leq I(B + A - \int (s')^3 p)$. Thus for $J = \log(I/q)$,

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

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 $\mathbb{E}[X_q h(X_q)] = 0$,*

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

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 $j b_j(X)^2$. 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 $(\mathbb{E}b_1)^2 \leq q\,\mathrm{Var}(b_0)/(1-q)$. The one-dimensional log-concave Poincaré inequality [3], applied in X, yields

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

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 $q > 0$,*

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

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

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

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 $\mathsf{H}(L, (p_Z + \phi)/2) \le C I(\mu)$. 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 $s_* = q w_*$ 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

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

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 $p_Z(z) \ge c_1 e^{-c_2 z^2}$ 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 $-\kappa \le S' \le 1$, the cubic terms satisfy

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

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

Finally $\|Q\|_{L^2(p_Z)} \le I(p_q)$ 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 $4a^2b^2 \le 1$, this proves the claimed estimate on this subinterval. In particular it covers $1/5 \le t \le 4/5$.

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:

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

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 $\mathsf{H}_\omega(r\phi, \phi)^2 \le C I(r\phi)$. Indeed for $f = \sqrt{r} - 1$, $\|xf\|_2 \le C(\|f\|_2 + \|f'\|_2)$ 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 $\int \omega T v \, d\gamma \le C \int \omega v \, d\gamma$. 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 $\sup_{n \ge 0} \sqrt{n+1}\,|m_{n+1} - m_n|$ 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

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

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 $\mathbb{E}[G r(Y)]$. 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,

$$\|J - b^3 A_3(q)e_3\|_{1,Z} = o(b^3)\sqrt{I(p_v)}.$$

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 $(d/r)^j e_j(R/r)$. 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 $r^j e_j(Z)$ 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 $\sqrt{2}\,d^j A_j(v) = b^j A_j(q)$ 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, $p_R \ge c\phi_{\sigma_1}$ 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:

$$\|\mathbb{E}[s_v(Y) \mid dY + sG]\|_2 \le C_s d^3\sqrt{I(p_v)}. \tag{89}$$

□

14.7 Exterior weights and the endpoints

Continue to assume $a \ge b$ and $b^2 \le 1/5$. 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 $b^2 \le 1/5 < 1/4$. 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 $j \le N$,

$$\mathbb{E}[Se_j(Z)] = (a^{j+1} + b^{j+1})A_j(q) + O_N(K).$$

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 $I(p_v) \le CK$ 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:

$$\|S\|_4^2 \le I(p_Z) + C\|S'\|_4^2 \le Cq\sqrt{K}.$$

The last step uses Lemma 14.2, $I(p_Z) \le CK$ and $K \le Cq^2$. In particular $\|SS'\|_2 \le Cq\sqrt{K}$, and the polynomial error above is $O_N(q\sqrt{K})$.

On any compact range $b_* \le b \le 1/\sqrt{5}$ use Lemma 14.7. The leading summands in the pairing $\mathcal{E}(S, J)$ are

$$(j-1)b^j(a^{j+1} + b^{j+1})A_j(q)^2 \ge 0.$$

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,

$$\mathcal{E}(S, J) \ge -o(b^3 K), \qquad \|J\|_{1,Z} \le Cb^3\sqrt{K}.$$

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 $Cqb^2 K$ 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

$$H = -\kappa I_n + M, \qquad M = \frac{q}{(1-q)^2}\operatorname{Cov}(X \mid X_q = x), \qquad -\kappa I_n \preceq H \preceq I_n.$$

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 $\{\lambda \ge K_n q\}$ move radially inwards by $\delta = \epsilon_n'/\sqrt{q}$ with $0 < \epsilon_n' \le \epsilon_n/4$. Choose K_n large and ϵ_n' small. The shifted point y has $\lambda(y) \ge q(K_n^{-1/2} + C\epsilon_n')^{-2} \ge 4q$, hence $\|H(y)\|_{\mathrm{HS}} \ge q$. 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

$$\Pr(\lambda \ge K_n q) \le C_n q^{-2} e^{-c_n/q} A(p_q).$$

On its complement $\|H\|_{\mathrm{HS}} \le C_n q$, and globally $\|H\|_{\mathrm{HS}} \le C_n$. This proves the first assertion.

For the OU evolution the identities are

$$\dot I = -2I - 2A, \qquad \dot A = -4A - 2B + 4\mathbb{E}\operatorname{tr}(H^3), \qquad \|L_p s\|_2^2 = B + A - \mathbb{E}\operatorname{tr}(H^3), \quad L_p = \Delta + \nabla \log p \cdot \nabla.$$

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 $s \leq 1/2$, $p_s \leq C_n$ 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 $D_{\mathrm{KL}}(p_s \| \phi_n) \leq C_{n,\ell} \mathsf{H}(p_s, \phi_n)^{2\ell}$ for every fixed $0 < \ell < 1$. OU entropy dissipation and monotonicity give $I(p_{3s/4}) \leq C D_{\mathrm{KL}}(p_s \| \phi_n)$. 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 $I(p_{16q}) \leq C_n q^2$ by the Fisher lemma and Poincaré, this gives

$$I(p_q) \leq C_{n,\ell} q^{2(1-\beta)} \mathsf{H}(p_{qw_*}, \phi_n)^{2\ell\beta}.$$

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 $A_\alpha(u) = \mathbb{E}[s_u(X_u)e_\alpha(X_u)]$. Coordinate-wise 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 $p \geq 2$. 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 $|y| \leq T$ 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 $|dy| \leq \sqrt{d}$. 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, $p_R \geq c_n\phi_{\sigma_1}$ 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 $\|\mathbb{E}[s_v(Y) \mid dY + sG]\|_2 \leq C_{n,s}d^3\sqrt{I(p_v)}$ for any other fixed $s > 0$ by the same proof. □

15.5 Closure at exterior weights

Keep $a \geq b$, $b^2 \leq 1/5$. 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 $\mathcal{E}(S,S) \geq c_n A(p_Z)$ 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,

$$\begin{aligned}\mathbb{E}[S_i e_\alpha(Z)] &= (a^{|\alpha|+1} + b^{|\alpha|+1})[A_\alpha(q)]_i + O_{n,\alpha}(K),\\ \mathcal{E}(S, \mathbf{e}_i e_\alpha) &= (|\alpha| - 1)\mathbb{E}[S_i e_\alpha(Z)] - \mathbb{E}[S_i S \cdot \nabla e_\alpha(Z)].\end{aligned}$$

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 $\|S\|_4^2 \leq I(p_Z) + C_n\|DS\|_4^2 \leq C_n q\sqrt{K}$ by Lemma 15.2 and $I(p_Z) \leq C_n K$. Consequently $\|HS\|_2 \leq C_n q\sqrt{K}$ 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

$$(|\alpha|-1)b^{|\alpha|}(a^{|\alpha|+1}+b^{|\alpha|+1})|A_\alpha(q)|^2 \geq 0.$$

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 $\|S\|_{1,Z} \leq C_n\sqrt{K}$. 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 $\mathcal{E}(S,J) \geq -o(b^3K)$ and $\|J\|_{1,Z} \leq C_n b^3\sqrt{K}$. 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

$$\|Q\|_2 \leq C_n b^3\sqrt{I(p_v)I(p_{v_h})} \leq C_n q b^3\sqrt{K}.$$

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.

16 Sharp high-signal endpoint asymptotics

16.1 Statement and context

For independent identically distributed X_1, X_2 and independent standard Gaussian G, set

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

when each input has Gaussian-relative density $1+g$. The normalization is mass one and covariance identity. We study the small bounded-relative neighborhood of the Gaussian among centrally even inputs.

Luo constructs near-Gaussian symmetric entropy counterexamples after smoothing and analyzes the Fourier resonance of periodic inputs without smoothing [7, Appendices C–E]. The periodic construction identifies surviving Fourier pairs at rational unequal weights and compares entropy values. We place the maximum of a two-frequency sum mode at a chosen weight and balance its amplitude against the other modes. This controls the second derivative at that weight with positive Gaussian noise. Abu-Ajamieh's scalar endpoint identity concerns the endpoint itself [2].

Theorem 16.1. *Fix $0 < t_* < 1/2$. If*

$$\frac{1-t_*}{1-t_*+t_*^2} < q \le 1, \tag{1}$$

there are smooth, even, variance-one, strongly log-concave scalar densities with Gaussian-relative perturbations arbitrarily small in sup norm for which $F''_{q,g}(t_*) > 0$. They can match the Gaussian moments through any prescribed finite order. Each fixed derivative of the relative perturbations tends uniformly to zero.

Corollary 16.2. *For $0 < q < 1$, let $a(q)$ be the supremum of the $a \in [0, 1/2]$ for which a positive dimension-independent bounded-relative radius ensures nonpositive curvature at every weight in $[0, a]$ for centrally even isotropic inputs. For $3/4 \le q < 1$,*

$$\frac{\sqrt{1-q}}{1+\sqrt{1-q}} \le a(q) \le \frac{\sqrt{(1-q)^2+4q(1-q)} - (1-q)}{2q}. \tag{2}$$

In particular $a(q)/\sqrt{1-q} \to 1$ as $q \uparrow 1$. At $q = 1$, no positive interval of fixed width has a common radius, even among smooth strongly log-concave scalar inputs.

16.2 Separating the Fourier exponents

Put $v = \sqrt{(1-t_*)/t_*} > 1$ and $D = 1 + v^2$. Write $r = q(1-t)$, $s = qt$. Given the scalar Gaussian output y, the mean of the two input normals is $(\sqrt{r}\, y, \sqrt{s}\, y)$, and their conditional covariance is $I - (\sqrt{r}, \sqrt{s})^{\mathsf{T}}(\sqrt{r}, \sqrt{s})$. For a product of cosines with input frequencies $a\xi, b\xi$, the sum mode has frequency $\omega_{a,b}(t)\xi$ and exponential factor $\exp(-E_{a,b}(t)\xi^2/2)$, with

$$\omega_{a,b}(t) = a\sqrt{r} + b\sqrt{s}, \qquad E_{a,b}(t) = a^2 + b^2 - \omega_{a,b}(t)^2.$$

This follows directly from the conditional Gaussian characteristic function. The difference mode of the same pair has a larger exponent.

At t_* let

$$A = (1-q)D, \qquad B = 2 - q(1 + 2v/D), \qquad C = 1 - qv^2/D.$$

These are the sum-mode exponents of $(v,1)$, $(1,1)$, and the least linear exponent of the low frequency, respectively. Equation (1) is exactly $A < C$. Choose a positive η satisfying

$$\max(0, A-B) < \eta < \min(v^2B - A,\ v^2C - A). \tag{3}$$

Such a choice exists. Indeed $DB > 2A$ because $2v/D < 1$ and $q > 0$. In particular $v^2B > A$ and $A - B < v^2B - A$. Also $v^2C > A$. For the other comparison, $B + v^2C - 2A$ is affine in q with positive slope $(v^4 + 3v^2 + 1 - 2v)/D$. At $q = v^2D/(v^4 + v^2 + 1)$, where $A = C$, its value is $2v^3(v-1)/(v^4 + v^2 + 1) > 0$. This proves (3).

Take the two-frequency function

$$f_\xi(x) = \cos(v\xi x) + e^{-\eta\xi^2/2}\cos(\xi x).$$

At t_* the conditional product has one dominant sum mode, with v in the larger load and one in the smaller. Its combined exponent is $A + \eta$. The two equal-frequency pair exponents are $B + 2\eta$ and v^2B, both larger by (3). Swapping the input frequencies in the mixed pair reduces the sum frequency since $r > s$ and $v > 1$, hence increases the exponent. Its difference modes also have higher exponents. The least exponents of the two linear terms are $C + \eta$ and v^2C, both strictly above $A + \eta$. The same inequalities hold when $q = 1$, where the conditional covariance is positive semidefinite.

16.3 Moment correction and the entropy sign

Choose a nonnegative even smooth compactly supported χ positive on an interval. For the even moments through the prescribed order, including mass and variance, use correcting functions $\chi(x)x^{2j}$. Their moment Gram matrix is positive definite. Every Gaussian moment of the two cosines is bounded by a polynomial in ξ times $\exp(-P\xi^2/2)$, where

$$P = \min(v^2, 1 + \eta) > A + \eta.$$

The bound for v^2 follows from $v^2C > A + \eta$ and $C < 1$, and $1 + \eta > A + \eta$ follows from $A < C < 1$. Solving the fixed linear moment system corrects f_ξ to a function $\widetilde{f}_\xi$ with exactly zero Gaussian moments of the required orders. The correction is of this smaller exponential size in all fixed derivative norms.

At the fixed interior weight, these estimates also bound the first two weight derivatives of all correction terms after taking the conditional product. To see this, use independent standard normals H, J and the conditional pair

$$Z_1 = \sqrt{1-t}\,(\sqrt{q}\,y + \sqrt{1-q}\,J) + \sqrt{t}\,H,$$
$$Z_2 = \sqrt{t}\,(\sqrt{q}\,y + \sqrt{1-q}\,J) - \sqrt{1-t}\,H.$$

On a compact interior weight interval, the derivatives of these affine arguments have Gaussian moments of every finite order. Derivatives of the cosine factors cost only polynomials in ξ, and one correction factor or its spatial derivatives costs the smaller exponential. This gives the same negligible bounds in Gaussian L^2, including for the linear correction.

Put $g_\xi = \epsilon_\xi \widetilde{f}_\xi$ with $\epsilon_\xi = e^{-\xi}$. The prescribed moments stay exact. The sup norm and all fixed positive derivative norms tend to zero, and $(1+g_\xi)\gamma_1$ is positive and strongly log-concave for large ξ, by differentiating the logarithm of its ratio. This amplitude is subexponential in ξ^2, so it preserves all strict exponent gaps.

Let $R(t,y)$ denote the Gaussian-relative output correction. At t_* its leading term, also for the first and second weight derivatives, comes from

$$M(t,y) = \frac{\epsilon_\xi^2}{2} \exp\left(-\frac{[E_{v,1}(t) + \eta]\xi^2}{2} \right) \cos(\omega_{v,1}(t)\xi y).$$

Every difference $R^{(j)} - M^{(j)}$, $0 \le j \le 2$, is exponentially smaller in L^2 than $\|M\|_2$ at t_*. Polynomial factors from weight derivatives do not affect the gaps in Section 16.2. The frequency $\omega = \omega_{v,1}$ attains a strict nondegenerate maximum at t_*. There $\omega' = E' = 0$, $M' = 0$, and $E'' = -2\omega\omega'' > 0$. Consequently

$$-\langle M'', M\rangle = \left(\tfrac{1}{2}E''\xi^2 + o(\xi^2)\right)\|M\|_2^2, \qquad \|M''\|_2 = O(\xi^2\|M\|_2).$$

The term from differentiating the frequency is negligible in the pairing, since the Gaussian squared norm of $\cos(\omega\xi\,\cdot)$ is $(1 + e^{-2\omega^2\xi^2})/2$.

The output variance is exactly one. Its entropy Hessian is therefore

$$F'' = -\langle R'', \log(1+R)\rangle - \int \frac{(R')^2}{1+R}\, d\gamma_1.$$

Conditioning bounds the ratio uniformly near one as $\xi \to \infty$. The difference of the first pairing from $-\langle M'', M\rangle$ is $o(\xi^2\|M\|_2^2)$, by the exponent gaps and $|\log(1+R) - R| \le C\|R\|_\infty |R|$. The last integral is of the same smaller order, since $M'(t_*) = 0$. The strictly positive leading coefficient proves Theorem 16.1.

Proof of Corollary 16.2. The unequal-weight curvature theorem of Section 13 gives the lower bound in (2). Rewriting (1) as $qt^2 > (1-q)(1-t)$ shows that any interval extending past the positive root of $qt^2 = (1-q)(1-t)$ contains a weight with arbitrarily small counterexamples. This is the upper bound in (2). For $\delta = 1 - q$ the lower and upper widths expand as $\sqrt{\delta} - \delta + O(\delta^{3/2})$ and $\sqrt{\delta} - \delta/2 + O(\delta^{3/2})$, respectively. They give the same leading constant. At $q = 1$ Theorem 16.1 applies at every unequal fixed interior weight, so no common positive endpoint interval survives.

16.4 Sharpness of a conditional product estimate

Let $B_{r,s}(g)$ denote the Gaussian conditional expectation of $g(Z_1)g(Z_2)$ given $\sqrt{r}\,Z_1 + \sqrt{s}\,Z_2 + \sqrt{1-r-s}\,G$. For $r \ge (\sqrt{5}-1)/2$ and $0 < s < 1 - r$, the bound

$$\|\mathcal{N} B_{r,s}(g)\|_2 \le C_{r,s}\|g\|_\infty \|(\mathcal{N}+1)\mathcal{T}_r g\|_2 \tag{4}$$

holds in every dimension with the same constant when $s \le r(1-r)$, by the binomial estimate of Section 13. Here $\mathcal{T}_r$ has Hermite multiplier $r^{n/2}$ and $\mathcal{N}$ has multiplier n. For $s > r(1-r)$, even the weaker estimate with $\|B_{r,s}(g)\|_2$ on the left fails for any finite constant in dimension one, within the same smooth moment-matched class.

Indeed, take $v = \sqrt{rs}/(1-r) > 1$ and use $\cos(v\xi x) + \cos(\xi x)$ before correcting its moments and scaling. The high–low sum mode has frequency $(v\sqrt{r} + \sqrt{s})\xi$ and exponent

$$E = (1-r)v^2 - 2\sqrt{rs}\,v + 1 - s = 1 - \frac{s}{1-r} < 1 - r.$$

Its frequency is distinct up to sign from all the others. The two equal-frequency sum modes are on either side of it, and the swapped sum is smaller since $r > s$. All difference frequencies have

modulus at most $v\sqrt{r}$. After grouping any other coincident modes, the Gaussian Gram matrix of the finitely many distinct cosine frequencies $c_j\xi$ has off-diagonal entries tending to zero, by the Gaussian characteristic function. Its limiting diagonal is positive, including any constant mode. Therefore $\|B_{r,s}(g)\|_2$ before correction and scaling is at least a constant times $\exp(-E\xi^2/2)$. Direct OU action on both cosines gives an upper bound $C(1+\xi^2)\exp(-(1-r)\xi^2/2)$ for the number-weighted norm on the right of (4).

The moment correction above now has exponent at least one, strictly exceeding E and $1-r$. Conditional contraction bounds its product errors, and its number norm controls the correction on the right. Thus the ratio diverges exponentially. Scaling preserves the ratio and supplies the same bounded smooth strongly log-concave witnesses. After correction the output product has mean $(\mathbb{E}g)^2 = 0$, so its number norm is at least its L^2 norm. This gives the stated exact boundary for (4) in the large-load range. The entropy curvature theorem uses the full Hessian, with its additional signed terms.

17 Eventual concavity and moment-matched counterexamples on the simplex

17.1 Statements and predecessors

Let X be isotropic in $\mathbb{R}^d$, meaning $\mathbb{E}X = 0$ and $\mathrm{Cov}(X) = I_d$. For independent copies X_i and independent standard normal G, put

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

Section 4 gives fixed-input eventual curvature in the scalar two-copy case with finite moments. Here the extension to the full simplex needs control of the Hessian even where its leading terms degenerate. Madiman, Nayar and Tkocz [8] used a fourth-cumulant entropy correction to obtain counterexamples to Schur concavity. The pinned Evidence Press candidate [5, Sections 7 and 9.3] adapted that correction to three copies and Gaussian noise, and asked what happens if the fourth moment matches the Gaussian. Section 7 already answers that question for fixed signal below one by a higher-order construction with weights approaching a face. We quantify the curvature crossover, include signal one, and obtain a fixed positive interior weight for each signal.

Write ϕ_d for the standard Gaussian density and He_α for product probabilists' Hermites. Suppose the first mismatch with the Gaussian is at total degree $r = 2k$, $k \geq 2$. Define the nonzero polynomial

$$P(x) = \sum_{|\alpha|=2k} \frac{\Delta_\alpha}{\alpha!}\mathrm{He}_\alpha(x), \qquad \Delta_\alpha = \mathbb{E}X^\alpha - \mathbb{E}G^\alpha.$$

For a simplex tangent v put

$$S_k(w) = \sum_i w_i^k, \qquad D_k(w,v) = \sum_i v_i^2 w_i^{k-2}, \qquad U_k(w,v) = \sum_i v_i w_i^{k-1},$$

with $0^0 = 1$ when $k = 2$.

Theorem 17.1. *Fix d, m and a centrally symmetric isotropic X having first mismatch at degree $2k$ and $\mathbb{E}|X|^{4k+2} < \infty$. As $q \downarrow 0$, uniformly in simplex weights and tangents,*

$$\partial_v^2 F_{X,m}(q,w) = -q^{2k}\|P\|^2_{L^2(\phi_d)}\left[k(k-1)S_k(w)D_k(w,v) + k^2 U_k(w,v)^2\right] + o(q^{2k})D_k(w,v).$$

Derivatives on faces are interpreted by continuous extension of the squared-gain derivatives. In particular for some $q_X, c_X > 0$, the Hessian is bounded above by $-c_X q^{2k} D_k(w, v)$ for $0 < q \le q_X$. The entropy is strictly concave on the entire simplex.

Every fixed centrally symmetric non-Gaussian isotropic log-concave input has a first mismatch and all moments. More generally Gaussian determinacy of all one-dimensional projections gives a first mismatch for every fixed non-Gaussian law with all moments.

Theorem 17.2. *Fix $m \ge 3$, $d \ge 1$, and even $r \ge 4$. For each $0 < q \le 1$ there are smooth strongly log-concave, centrally symmetric isotropic inputs arbitrarily close to the Gaussian in every fixed bounded relative C^s norm, matching all Gaussian moments of total degree $< r$, with first mismatch at degree r, for which $F_{X,m}(q, \cdot)$ has strictly positive curvature at an interior weight. In dimension one and for three copies, one can use the segment $(1-\delta, \delta(1-t), \delta t)$. For any even $L > r$, examples with relative sup norm comparable to ϵ can be chosen so that the split curvature is positive on this full segment for every*

$$0 < \delta \le c\, \epsilon^{4/(L-r)} q^{2L/(L-r)}.$$

The constants are independent of $q \in (0, 1]$.

17.2 The fixed-law theorem

Use a real gain vector $h = (h_1, \dots, h_m)$ in a small Euclidean ball. The gain output is $\sum_i h_i X_i + (1 - |h|^2)^{1/2} G$. Let $\mathcal{D}(h)$ be its relative entropy to ϕ_d. Squared gains will be denoted by $u_i = h_i^2$. We give the finite-jet argument, including the face identities that yield the weighted error.

For any derivative of total gain order j, the density derivative divided by the density is a conditional expectation of a polynomial in $(X_1, \dots, X_m, G)$. On a fixed compact gain ball, its coefficients are bounded, its latent degree is at most j, and its residual Gaussian degree at most $2j$. Indeed the kernel can be written as an isotropic Gaussian with mean $\sum_i h_i X_i$ and variance $1 - |h|^2$. Each derivative adds at most one latent degree and two Gaussian degrees, by differentiating the mean, variance and then the standardized residual in the preceding polynomial. On bounded spatial sets all these kernel derivatives can be passed through the expectation using the finite moments.

The density has a uniform quadratic logarithmic envelope on the compact gain ball, by a Gaussian upper bound and a lower bound obtained by restricting the inputs to a bounded positive-probability set. Formal entropy derivatives of order j consist of a density derivative times this logarithm (plus a bounded part) and products of normalized density derivatives with orders summing to j, integrated against the density. Conditional Jensen and weighted Young bound each product in absolute value by conditional polynomial moments of latent degree at most j and Gaussian degree at most $2j$. The log factor costs two further degrees in each. Since the input moment $4k+2$ is finite, the derivative tails vanish uniformly on this gain ball for $j \le 4k$: if the output norm is large then the sum of the latent and Gaussian norms is large. Apply the derivatives first with spatial cutoffs, then remove them using these uniform bounds. This extends the finite-moment kernel argument of Sections 3–4 to the present m, d, and proves $\mathcal{D} \in C^{4k}$.

By central symmetry the gain entropy is even in each coordinate. The elementary even-square formula

$$\frac{d^j}{du^j} E(\sqrt{u}) = \frac{1}{2^{2j-1}(j-1)!} \int_0^1 (1 - s^2)^{j-1} E^{(2j)}(s\sqrt{u})\, ds \qquad (j \ge 1)$$

follows by the fundamental theorem of calculus and induction by parts. Apply it coordinatewise to obtain a C^{2k} squared-gain entropy $D(u) = \mathcal{D}(\sqrt{u_1}, \dots, \sqrt{u_m})$ with continuous derivatives up to the faces.

The gain density jet has no nonconstant term below total order $2k$, and its term of order $2k$ is

$$\phi_d(x)P(x)\sum_i h_i^{2k}.$$

This follows by Taylor differentiation of the product of the characteristic functions of the copies after factoring out the standard Gaussian characteristic function. In each factor the first nonconstant term is the degree-$2k$ Gaussian moment discrepancy, and Gaussian differentiation gives the stated Hermite polynomial. The entropy jet through order $4k$ has only the square of this leading density term, since mass kills the linear term and the first cubic term has degree $6k$. Thus

$$D(u) = \tfrac{1}{2}\|P\|_2^2\left(\sum_i u_i^k\right)^2 + R(u),$$

where R has zero squared-gain jet through degree $2k$. This is a statement about finite Taylor jets, with the differentiated entropy integral justified above.

There is additional vanishing away from the origin. On each face $u_i = 0$ the traces of $\partial_i^j D$ vanish for $1 \le j < k$, with the other gains fixed. Indeed, replace X_i by a standard Gaussian. Its gain is then immaterial, because its variance is absorbed by the residual Gaussian. The density derivatives with respect to h_i at zero through order $2k-1$ agree with this replacement by moment matching. All associated entropy derivatives agree too. The traces of R satisfy the same vanishing identities after positive normal derivatives of orders $< k$, since the displayed leading polynomial also has those zero traces. Their tangential derivatives vanish whenever defined.

Let $Q = \sum_i u_i$. The zero jet at the origin gives $\partial_i^k R = o(Q^k)$ and $\partial_i^k\partial_j^k R = o(1)$. For $k > 2$ integrate the first bound $k-2$ times in u_i, using the zero traces, to get $\partial_{ii}R = u_i^{k-2}o(Q^k)$. For $k = 2$ the same bound holds directly. For $i \ne j$ integrate the second bound $k-1$ times in each coordinate to obtain $\partial_{ij}R = u_i^{k-1}u_j^{k-1}o(1)$. All remainders are uniform in the positive orthant near the origin, since the integrated coordinate segments stay there with total gain at most Q. At $u = qw$, multiply by q^2 for the two weight derivatives and sum against v_iv_j. The mixed terms are controlled using

$$\left(\sum_i |v_i| w_i^{k-1}\right)^2 \le D_k(w,v)S_k(w).$$

This gives the claimed $o(q^{2k})D_k(w,v)$. The covariance of the gain output is I_d, so entropy is $h(\phi_d) - D$. The Hessian of the leading square gives the formula in the theorem, and $S_k(w) \ge m^{1-k}$ gives its strict bound along every nonconstant simplex segment. Indeed in the relative interior of a segment every coordinate with nonzero directional component has positive weight, so $D_k > 0$ there.

The proof also gives uniform restoration for a family with a common bound on $\mathbb{E}|X|^{4k+4}$, exact matching below degree $2k$, and a positive common lower bound on $\|P\|_2$. The extra two input moments give bounded derivatives of squared-gain order $2k+1$, so the little-oh error is bounded uniformly by $Cq^{2k+1}D_k$ by the same trace integrations.

17.3 A first-unmatched-moment obstruction

We work first on the line. For even $r \ge 4$ and a smooth near-Gaussian density p, put

$$B_r(p) = \int p^{(r)} \log p.$$

Let $p = \phi(1+u)$, where $\|u\|_\infty \le 1/2$ and u is smooth with finite Gaussian Sobolev norms. Since $\log\phi$ is quadratic and r is even, integration by parts and the logarithm expansion give

$$B_r(p) = \langle \mathrm{He}_r, u\rangle_\gamma + Q_r(u) + \mathrm{Err}_r(u), \tag{95}$$

$$Q_r(u) = \langle u^{(r)}, u\rangle_\gamma - \tfrac{1}{2}\langle \mathrm{He}_r, u^2\rangle_\gamma. \tag{96}$$

For a sufficiently large fixed integer $s = s(r)$,

$$|\,\mathrm{Err}_r(u)| \le C_r\|u\|_\infty \|(1+\mathcal{N})^s u\|_2^2,$$

where $\mathcal{N}$ is Gaussian Hermite degree. Indeed $\phi^{-1}(\phi u)^{(r)}$ is a finite combination of derivatives of u times polynomials, bounded from this Gaussian Sobolev norm to L^2 by the Gaussian raising and lowering identities. The $\phi^{(r)}$ part uses the cubic error in $\log(1+u)$, and the $(\phi u)^{(r)}$ part uses the quadratic error there. This proves the bound by Cauchy–Schwarz and the same polynomial operators. It also proves continuity of Q_r in this norm.

Use normalized Hermites $e_j = \mathrm{He}_j/\sqrt{j!}$. For any even $L > r$ the Hermite product formula gives

$$Q_r(e_L) = -\tfrac{1}{2}\langle \mathrm{He}_r, e_L^2\rangle_\gamma < 0.$$

The derivative pairing vanishes by orthogonality, and the inner product in the second term is $r!(L - r/2)!\binom{L}{r/2}^2/L! > 0$. Choose a compactly supported even smooth f very close to e_L in the Gaussian Sobolev norm, with no coefficients below degree L. This can be done by cutting off e_L and restoring the finitely many moments with fixed even compactly supported smooth functions. For instance the functions χe_{2j}, with a nonnegative smooth bump χ positive on an interval, have a positive definite moment matrix since polynomials are linearly independent there. The correction coefficients tend to zero as the cutoff increases. The Gaussian Sobolev convergence of the cutoffs follows by applying powers of the number operator, a differential operator with polynomial coefficients, to the tails. Similarly choose an even compactly supported smooth h with no coefficients below L except for degree r, with $\mu_r = \langle \mathrm{He}_r, h\rangle > 0$. Let $\mu_L = \langle \mathrm{He}_L, f\rangle \ne 0$. If $0 < s_0 \le 1$, write $T_{\sqrt{s_0}}$ for the OU operator with multiplier $s_0^{j/2}$ at degree j. The convergence of f can be made close enough that

$$Q_r(T_{\sqrt{s_0}} f) \le -c\, s_0^L \qquad (0 < s_0 \le 1).$$

In fact $s_0^{-L/2} T_{\sqrt{s_0}}(f - e_L)$ has norm bounded by the original approximation error because all its degrees are at least L.

For $0 < q \le 1$ use the input relative perturbation

$$g = \epsilon f + \lambda \epsilon^2 q^{L - r/2} h$$

with a fixed sufficiently small $\lambda > 0$. For small $\epsilon > 0$ the density $\phi(1+g)$ is positive and strongly log-concave, since the first two derivatives of g are uniformly small. It matches all lower moments and has first mismatch at r. Its smoothed p_q has $\kappa_r(p_q) = \lambda\epsilon^2 q^L \mu_r > 0$. Here and below κ_r equals the r-th moment difference because all lower moments match. For p_{heavy} the smoothing of the same input at $q(1-\delta)$, (95) yields

$$B_r(p_{\mathrm{heavy}}) \le -c\epsilon^2 q^L \qquad (0 \le \delta \le \delta_0) \tag{97}$$

for fixed small $\delta_0 > 0$, uniformly in q. The negative contribution from f is of size $\epsilon^2[q(1-\delta)]^L$, while the positive linear term from h is of size $\lambda\epsilon^2 q^L (1-\delta)^{r/2}$. Choose λ small depending on the fixed δ_0, L. The h contribution to the smoothed Gaussian Sobolev norm is $O(\epsilon^2 q^L)$, so mixed quadratic terms and the remainder in (95) are $O(\epsilon^3 q^L)$. This proves (97).

17.3.1 The light weights

For three copies put $w = (1-\delta, \delta(1-t), \delta t)$, $P_j(t) = (1-t)^{j/2} + t^{j/2}$, and denote by $1+R$ the relative output density and by $1+W$ the relative density p_{heavy}/ϕ. In particular W is independent of the split parameter t. We justify the expansion uniformly through $t = 0, 1$, including at $q = 1$.

For normalized Hermites the output coefficient at total degree N is

$$\sum_{j_0+j_1+j_2=N} \binom{N}{j_0, j_1, j_2}^{1/2} \prod_{i=0}^{2} w_i^{j_i/2} \, \widehat{(1+T_{\sqrt{q}} g)}_{j_i}.$$

This follows by conditioning independent Gaussian inputs on their normalized linear combination and using the Hermite generating functions. Each nonconstant index is r or an even degree $\geq L$. Group terms by which light indices are occupied and which of these two classes they belong to. For an occupied light index strip r or L powers of the linear gain according to its class. If the total number of stripped powers is J, the resulting terms have a factor $\delta^{J/2}$. Up to two split derivatives reduce only powers of $t, 1-t$; the stripped factors leave all those powers nonnegative since each stripped index is at least four. The differentiation costs at most $(N+1)^2$. The ratio of the original multinomial coefficient square root to the degree-$(N-J)$ one is bounded by $(N+1)^{J/2}$, by the factorial formula. The remaining weights are square roots of multinomial probabilities at degree $N-J$. Cauchy–Schwarz bounds the squared sum of coefficients by the corresponding sum of squared products times this polynomial in N. Summing over N is bounded by the product of fixed Gaussian Sobolev norms, since $N+1 \leq \prod_{i=0}^{2}(j_i+1)$. It follows also that the coefficient series and its first two split derivatives converge uniformly in Gaussian L^2 on the whole split interval.

Specifically put

$$A = \epsilon^2 q^L \delta^{r/2}, \qquad B = \epsilon q^{L/2} \delta^{L/2}.$$

The norms of $R - W, R_t, R_{tt}$ are $O(A+B)$. Removing the singleton degree-r light term from R_{tt} leaves an $O(B + A^2)$ norm. These estimates follow from the preceding class bound: one occupied r slot has norm $O(A)$, one high slot has norm $O(B)$, and two occupied light slots multiply their respective costs. The heavy slot has uniformly bounded Sobolev norms. This argument retains its bounds without any smoothing reserve, because compact smooth g has every fixed Gaussian Sobolev norm finite. Conditioning also gives uniform positive upper and lower bounds for $1+R$ and $1+W$, including at $q = 1$.

Differentiate entropy in $L^2(\gamma)$ using these bounds. Replacing $\log(1+R)$ by $\log(1+W)$ in the R_{tt} pairing, and paying the $\int R_t^2/(1+R)\, d\gamma$ term, costs $O((A+B)^2)$. Also $\|\log(1+W)\|_2 = O(\epsilon q^{L/2})$. The singleton r-term is $\delta^{r/2}\kappa_r(p_q) P_r''(t)\, p_{\mathrm{heavy}}^{(r)}/(r!\phi)$. This identity follows by multiplying the Hermite coefficients above with the light index fixed at r, or by the Gaussian identity $e_j \phi = (-1)^j \phi^{(j)}/\sqrt{j!}$. Its pairing with $\log(1+W)$ is the same as its pairing with $\log p_{\mathrm{heavy}}$, since the Gaussian logarithm is quadratic. Consequently the split curvature is

$$-\frac{\delta^{r/2}\kappa_r(p_q)}{r!} P_r''(t) B_r(p_{\mathrm{heavy}}) + O\Big(\epsilon q^{L/2}(B + A^2) + (A+B)^2\Big). \tag{98}$$

In particular its leading term is positive and at least $c\epsilon^4 q^{2L} \delta^{r/2}$, since P_r'' is bounded below positively. The claimed upper bound on δ, with its constant small, makes the high-light error $\epsilon q^{L/2} B$ smaller than a fixed fraction of this. In that range $B \leq CA$, and the remaining errors have the further small factor $\delta^{r/2}$. This proves strict convexity on the full split segment in Theorem 17.2.

For more copies take the remaining weights initially zero, then increase them by continuity at any fixed positive δ and split $0 < t < 1$. The same Hermite coefficient bounds, with finitely many

additional factors and derivatives only on the two light slots, justify curvature continuity, including at $q = 1$. For $d > 1$ take an independent product with $d - 1$ standard normals. This preserves the moment matching, strong log-concavity and the relative C^s closeness, and only adds a constant to entropy.

17.3.2 The boundary limit and the crossover

The same coefficient proof gives a general boundary test for a bounded smooth even relative input with finite Gaussian Sobolev norms. If its first mismatch is at degree r, then for each fixed $q > 0$

$$\delta^{-r/2}\partial_t^2 F_{X,3}(q, (1-\delta, \delta(1-t), \delta t)) \;\longrightarrow\; -\frac{\kappa_r(p_q)}{r!} B_r(p_q) P_r''(t)$$

uniformly on $0 \le t \le 1$, where the derivative is along the indicated split. To see this, strip $r + 2$ in all light degrees above r. They contribute $O(\delta^{(r+2)/2})$. The two occupied light slots and the entropy square are $O(\delta^r)$, and $B_r(p_{\text{heavy}}) \to B_r(p_q)$ by the Gaussian Sobolev and bounded-relative continuity. Thus $\kappa_r(p_q)B_r(p_q) \ge 0$ is necessary for local concavity along all such splits. The fourth-order criterion is in [5, Proposition 7.2]. Section 7 uses the higher-order formula near balanced splits at positive noise. The limit here is uniform on the closed split interval and includes signal one.

For the fixed shapes f, h above, one can also identify the change of sign at the stated amplitude scale. Fix $q > 0$, let

$$C_q = -Q_r(T_{\sqrt{q}} f) - \lambda q^L \mu_r > 0, \qquad \delta = \xi\, \epsilon^{4/(L-r)}.$$

For ξ in a compact subinterval of $(0, \infty)$, the following limit is uniform in the split:

$$\frac{\partial_t^2 F_{X,3}}{\epsilon^4 \delta^{r/2}} \longrightarrow \frac{\lambda q^L \mu_r C_q}{r!} P_r''(t) - \frac{q^L \mu_L^2}{L!} \xi^{(L-r)/2} P_L''(t).$$

Indeed isolate the L-singleton as well as the r-singleton in the expansion of R_{tt}. The higher even degrees now strip $L + 2$. The L-singleton uses the Hermite moment $\mathbb{E}[\mathrm{He}_L(X_q)]$ in place of a cumulant. This moment is $\epsilon q^{L/2}\mu_L + O_q(\epsilon^2)$; the same expression, with additional $O_q(\epsilon\delta)$, is $B_L(p_{\text{heavy}})$ by (95). Meanwhile $B_r(p_{\text{heavy}})/\epsilon^2 \to -C_q$. The remaining estimates in (98) are little-oh at the crossover scale after isolating that singleton. Thus this scale gives positive curvature for small fixed ξ and negative curvature for large fixed ξ, for this family.

17.4 Counterexamples at a fixed positive weight

The preceding examples have a fixed high degree L. Allowing that degree to increase gives a counterexample on a fixed interior split at each signal, as follows.

Theorem 17.3. *Fix an even $r \ge 4$. There is $c_r > 0$ such that for each $q \in (0, 1]$ one can fix $\delta = c_r q^2 > 0$ and find a sequence of smooth strongly log-concave, symmetric scalar inputs converging to the Gaussian in every fixed bounded relative C^s norm, matching its moments below degree r, with first mismatch at r, and having strictly positive split curvature at $(1-\delta, \delta/2, \delta/2)$. In fact the split curvature is positive on the full segment for this δ.*

Proof. Here L tends to infinity through even integers $> r$. We first specify quantitative choices of the test shapes. Let χ_L be an even smooth cutoff equal to one on $[-C\sqrt{L}, C\sqrt{L}]$ and zero beyond

$[-2C\sqrt{L}, 2C\sqrt{L}]$, obtained by rescaling a fixed cutoff taking values in $[0,1]$. Choose C sufficiently large. For every fixed Gaussian number Sobolev norm,

$$\|(1+\mathcal{N})^s[(1-\chi_L)e_j]\|_2 \le C_s e^{-cL} \quad (0 \le j \le L).$$

Indeed powers of $\mathcal{N}$ act by a fixed number of derivatives and polynomial multiplications, which on Hermites cost fixed powers of $L+1$ by the raising and lowering identities. All these terms are supported on the Gaussian tail beyond $C\sqrt{L}$. Their L^2 norms there follow from Cauchy–Schwarz, the Gaussian tail and $\|e_j\|_4 \le 3^{j/2}$. Derivatives of the cutoff are bounded by powers of $L^{-1/2}$. Taking C large absorbs the exponential L^4 bound, uniformly for each fixed order as $L \to \infty$.

The truncated Gram matrix $(\langle \chi_L e_i, e_j\rangle_\gamma)_{i,j<L,\ i,j \text{ even}}$ differs from the identity in operator norm by $O(Le^{-cL})$, so it is invertible for large L. Correct $\chi_L e_L$ using its columns to obtain f_L with degrees at least L. Correct $\chi_L e_r$ to obtain h_L whose coefficients below L are exactly those of e_r. The correction coefficients have exponentially small absolute sums. Thus $f_L - e_L$ and $h_L - e_r$ are exponentially small in every fixed Gaussian number Sobolev norm. Furthermore for some C_1 independent of the derivative order,

$$\|f_L\|_{C_b^s} + \|h_L\|_{C_b^s} \le C_s(L+1)^{C_s} e^{C_1 L}.$$

Indeed on $|x| \le 2C\sqrt{L}$ the Hermite generating function and the Cauchy bound at radius $\sqrt{j}$ give an exponential bound $e^{C_1 L}$ for every normalized degree $j \le L$. Fixed derivatives use the lowering identities and derivatives of the cutoff, costing polynomial factors.

Use a smoothing reserve $q' = 1/2$, and put $Q = qq'$, $\delta_0 = cQ^2$ for a small constant $c > 0$ to be chosen. With $A > C_1 + 1$ fixed let

$$\epsilon_L = e^{-AL}, \qquad \lambda_L = \lambda(1-\delta_0)^{L-r/2}L^{r/2}, \qquad \widetilde{g}_L = \epsilon_L f_L + \lambda_L \epsilon_L^2 Q^{L-r/2} h_L.$$

The input relative perturbation is $g_L = T_{\sqrt{q'}}\widetilde{g}_L$. For large L it is arbitrarily small in every fixed bounded C^s norm by the preceding estimates and OU derivative contraction. In particular the input is positive and strongly log-concave. All its moments below r match, and the mismatch at r is strictly positive. At the final smoothing signal q, its degree-r Hermite coefficient is $\lambda_L \epsilon_L^2 Q^L$.

Uniformly in $0 \le \delta \le \delta_0$, the heavy density satisfies

$$B_r(p_{\text{heavy}}) \le -c_r' \, \epsilon_L^2 [Q(1-\delta)]^L L^{r/2}.$$

To verify this, normalize the OU image of f_L at signal $s = Q(1-\delta)$ by $s^{-L/2}$. It stays exponentially close to e_L in every fixed Gaussian Sobolev norm, by its degree cutoff. The explicit product coefficient in (95) has magnitude comparable to $L^{r/2}$. The positive linear term from h_L, relative to the displayed negative quantity, costs at most a constant times $\lambda((1-\delta_0)/(1-\delta))^{L-r/2} \le C\lambda$. The other quadratic terms from h_L are smaller by an exponential in L, using $\lambda_L \le \lambda L^{r/2}$ and the Gaussian Sobolev bounds at the corresponding signal. The cubic error in (95) is bounded by a fixed polynomial in L times this main scale times $\|g_L\|_\infty$, which is exponentially small. Choose $\lambda > 0$ small.

We give the uniform error comparison for the light weights, because the constants in the preceding fixed-shape calculation need not be uniform in L. Compute the output from $\widetilde{g}_L$ at effective weights $Q(1-\delta), Q\delta(1-t), Q\delta t$, with Gaussian noise supplying the remainder. For terms with at least one high light degree, replace both reference light weights in the Gaussian multinomial bound by $c_1 = 1/8$ and keep the heavy effective weight. The reference sum is at most $3/4$; the missing probability can be assigned to a Gaussian index of degree zero. Each high light degree contributes

a factor at most $(Q\delta/c_1)^{L/2}$, for $Q\delta < c_1$, and differentiation preserves this factor, because the remaining split powers after at most two derivatives are nonnegative. Cauchy–Schwarz as before now costs only a fixed polynomial in L, from the $(N+1)^2$ derivative factors and the Sobolev norms. Thus the Gaussian L^2 norm of the high-light part, through two split derivatives, is bounded by

$$\mathrm{poly}(L)\,\epsilon_L(Q\delta/c_1)^{L/2}.$$

Pairing its second derivative with the heavy logarithm costs at most $\mathrm{poly}(L)\epsilon_L^2 Q^L(\delta/c_1)^{L/2}$, since the heavy logarithm has L^2 norm at most $\mathrm{poly}(L)\epsilon_L Q^{L/2}$. The high part's squared norm is bounded by the same error because $\delta_0 < c_1$.

For the singleton r and double r terms, the shift by r or $2r$ of the heavy coefficient series costs a constant depending only on r times a fixed Gaussian Sobolev norm of the heavy density ratio. These norms are bounded uniformly in L. Writing $M_r = \kappa_r(p_q)$, the two light norms and the contribution from changing the logarithm and the entropy square therefore cost, apart from high errors, at most $C_r M_r^2 \delta^r$ after the leading singleton is removed. Here we also used boundedness of the heavy logarithm. Compared to the magnitude of the leading singleton, this costs a relative $C_r(M_r/|B_r(p_{\mathrm{heavy}})|)\delta^{r/2}$. The choices above bound this by $C_r\lambda\delta^{r/2}$. Mixed low and high square errors are bounded by the sum of the two squares.

The leading positive curvature is bounded below by a constant times $\epsilon_L^4 Q^{2L}(1-\delta_0)^{2L}L^r\delta^{r/2}$. For $0 < \delta \le \delta_0$, the high error divided by this is at most

$$\mathrm{poly}(L)\,\delta_0^{-r/2}\left[\frac{e^{2A}(\delta_0/c_1)^{1/2}}{Q(1-\delta_0)^2}\right]^L.$$

Choose $c > 0$ sufficiently small in $\delta_0 = cQ^2$ so the bracket is below $1/2$ for every $0 < Q \le 1/2$, and the relative low error is below $1/4$. For each fixed $q > 0$ the high error tends to zero with L, uniformly in the split and $0 < \delta \le \delta_0$. This proves the theorem with $c_r = c/4$. □

18 The sharp trimming scale for multicopy concavity

18.1 Statements

Let X_i be independent copies of a random vector with density $(1+g)\gamma_d$, where γ_d is standard Gaussian measure in dimension d. Assume g is bounded, real and centrally even, and the density has mass one and covariance I_d. For simplex weights $w = (w_i)_{i=1}^m$ define

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

where G is independent standard normal and h is differential entropy. Write $W = \max_i w_i$ and $w_- = \min_i w_i$.

Theorem 18.1. *For each $m \ge 3$ there exist $c_m, \rho_m > 0$ depending only on m such that in every dimension the Hessian of $F_{q,g}$ is strictly negative on all nonzero simplex tangents, unless $g = 0$, whenever*

$$\|g\|_\infty \le \rho_m, \qquad 0 < q\sqrt{W/w_-} \le c_m.$$

In particular the same radius works throughout the trimmed simplex $w_- \ge C_m q^2$ for a sufficiently large constant C_m, independent of dimension and signal.

Theorem 18.2. *Fix $m \geq 3$ and an even $r \geq 4$. There is a constant $c'_{m,r} > 0$ such that, for every $q \in (0,1]$, one can fix $\delta = c'_{m,r} q^2$ and the interior weight*

$$(1-\delta, \delta/(m-1), \ldots, \delta/(m-1)).$$

At this same weight there are smooth strongly log-concave centrally symmetric isotropic inputs arbitrarily close to the Gaussian in every fixed bounded relative C^k norm, matching its moments through total degree $r-1$, whose Hessian is positive definite on tangents that only redistribute the light weights. This holds in every dimension.

Thus for each $m \geq 3$, as $q \downarrow 0$, the infimum trim for the existence of a Gaussian-relative concavity neighborhood is of order q^2; for necessity one may fix $r = 4$. The obstructions have their smallest weight at a constant times q^2, so they lie strictly inside any trimmed region with half that minimum requirement. The pinned Evidence Press candidate [5, Section 9.3] asked for local concavity on fixed interior regions. The earlier sufficient range in Section 9 was of order $\eta^{3/2}$ in the retained signal for a fixed trim η. Theorems 18.1 and 18.2 give the matching order $\eta^{1/2}$ for sufficiently small η. Section 7 and the exact first-unmatched-moment construction of Section 17 supply the near-face precedent. We use the fixed three-copy construction of Section 17 for the obstruction and extend it to arbitrary copy counts.

18.2 Conditional products and the linear energy

Let $\mathcal{T}_a = T_{\sqrt{a}}$ be the OU operator with Gaussian degree-n multiplier $a^{n/2}$, and let $\mathcal{N}$ be total Hermite degree. All input degrees of g are even and at least four. For $H \subseteq \{1, \ldots, m\}$ let B_H be the conditional expectation of $\prod_{i \in H} g(G_i)$ given $\sqrt{q}\sum_i \sqrt{w_i} G_i + \sqrt{1-q}G$, with independent standard normal inputs. The output density ratio is

$$1 + R = 1 + L + M, \qquad L = \sum_i \mathcal{T}_{q w_i} g, \quad M = \sum_{|H| \geq 2} B_H.$$

Choose the coefficient convention in which Hermites are normalized in Gaussian L^2. The term in B_H with input indices α_i has output index $\sum_i \alpha_i$, and coefficient

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

This follows by taking Gaussian conditional expectations in the Hermite generating function. In particular, for any $c_i > 0$ with $\sum_{i \in H} q w_i / c_i = 1$, the expression equals the conditional projection of the tensor $\bigotimes_{i \in H} \mathcal{T}_{c_i} g$ onto the normal combination with weights $q w_i / c_i$. The projection contracts L^2 and preserves the total input degree as the output degree.

We will use the following elementary tensor estimate, also used in Section 12. If $j = |H|$, $0 < c_i \leq b_0 < 1$, and $\prod_i c_i \leq a/2$, then for any fixed s the L^2 norm of the tensor after multiplying its degrees by a scalar of absolute value bounded by $(1 + \sum_i n_i)^s$ is at most

$$C_{j,s,b_0} \|g\|_\infty^{j-1} \|\mathcal{T}_a g\|_2.$$

Indeed set $k = \min_i n_i$ and $\ell = \sum_i n_i - jk$. The squared degree factor is bounded by a constant times a^k, since

$$(1 + jk + \ell)^{2s} \prod_i c_i^{n_i} \leq (1 + jk + \ell)^{2s} 2^{-k} a^k b_0^\ell.$$

Sum over the squared Hermite coefficients, taking each possible index of the minimum in turn, and use $\|g\|_2 \le \|g\|_\infty$ on the other factors. The same bound holds after any such conditional projection. All constants are independent of the spatial dimension.

Fix an interior weight and a tangent v. Let $u_i = v_i/w_i$, $\rho = \|g\|_\infty$ and

$$a_i = q\sqrt{W w_i}, \quad a_- = q\sqrt{W w_-}, \qquad D_i = \|\mathcal{N}\mathcal{T}_{a_i} g\|_2, \quad D_- = \|\mathcal{N}\mathcal{T}_{a_-} g\|_2.$$

Use primes for derivatives along v. At every Gaussian degree $n = 2k \ge 4$, the linear multipliers for L and L'' are

$$q^k \sum_i w_i^k, \qquad q^k k(k-1) \sum_i u_i^2 w_i^k.$$

It follows that the linear energy satisfies

$$E = \langle L'', L\rangle \;\ge\; c \sum_i u_i^2 D_i^2. \tag{99}$$

Indeed the first multiplier contains the summand $q^k W^k$, and $k(k-1) \ge n^2/8$ for $k \ge 2$.

From now on assume $q\sqrt{W/w_-} \le c_m$, with c_m small depending only on m. In particular $3q < 1$ and

$$9q^2/a_- \le 9mc_m.$$

Using the conditional estimate with the common levels $q\sum_{i\in H} w_i \le q$ for each subset gives

$$\|M\|_2 + \|\mathcal{N}M\|_2 \le C_m \rho D_-. \tag{100}$$

Here the product of levels is at most $q^2 < a_-/2$ if c_m is sufficiently small.

18.3 The Hessian estimates

For a subset H and $i, j \in H$ define $D_{ij}B_H = w_i w_j \partial_{ij} B_H$, using the coefficient formula to differentiate at w. Its multiplier on the input tensor is bounded by a constant times $(1 + \sum_{\ell\in H} n_\ell)^2$, with no inverse weight factor. For $i = j$ this uses $w_i^2 \partial_{ii} w_i^{n_i/2} = (n_i/2)(n_i/2 - 1) w_i^{n_i/2}$. The second derivative along v sums these normalized derivatives with coefficients $u_i u_j$.

The most asymmetric pairing is with L, whose largest signal can exceed a_-. Move each self-adjoint singleton smoothing onto the mixed derivative:

$$\langle D_{ij}B_H, \mathcal{T}_{qw_\ell} g\rangle = \langle \mathcal{T}_{qw_\ell} D_{ij} B_H, g\rangle.$$

In the right-hand tensor, the common level is now $b = q^2 w_\ell \sum_{i\in H} w_i \le q^2$. Indeed the added OU smoothing acts on total degree, so it multiplies every input level by qw_ℓ. Since $b \le a_-/2$, one can pay the normalized derivative cost by the spare attenuation $(b/a_-)^{\sum_i n_i/2}$, using $\sup_N (1+N)^2\, 2^{-N/2} < \infty$. Conditional contraction gives

$$\|\mathcal{T}_{qw_\ell} D_{ij} B_H\|_2 \le C_m \|\mathcal{T}_{a_-} g\|_2^{|H|}.$$

Pairing with g, and using $\rho \le 1$, bounds the original pairing by $C_m \rho D_-^2$. Since $\sum_{ij} |u_i u_j| \le m \sum_i u_i^2$,

$$|\langle M'', L\rangle| \le C_m \rho D_-^2 \sum_i u_i^2 \le C_m \rho E.$$

Without this added smoothing, the common-level tensor estimate used in (100) also controls $\|D_{ij}B_H\|_2 \le C_m \rho D_-$ by taking the degree cost $s = 2$. Thus for every V with $\|V\|_2 \le C_m \rho D_-$,

$$|\langle M'', V\rangle| \le C_m \rho^2 E.$$

In the linear derivative, the i-th component L_i'' satisfies

$$\|\mathcal{N}^{-1}L_i''\|_2 \le Cu_i^2\|\mathcal{N}\mathcal{T}_{qw_i}g\|_2 \le Cu_i^2D_i.$$

The inverse is zero on Gaussian constants. Together with (100), this controls $\langle L'', M\rangle$ by $C_m\rho E$.

We next control the logarithmic error $\psi(R) = \log(1+R) - R$. The density ratio is between $(1-\rho)^m$ and $(1+\rho)^m$, by conditional expectation. In a fixed small relative neighborhood this gives

$$|\psi(R)| \le CR^2, \qquad |\psi'(R)| \le C|R|, \qquad |\psi''(R)| \le C.$$

Write R_+ for the output perturbation at signal $3q$. Then $R = T_{1/\sqrt{3}}R_+$ by a common Gaussian smoothing of the output. For every nonempty subset in R_+, put $\mathcal{T}_{3q}g$ in its tensor and project to the corresponding output Gaussian sum. This is equivalent by the coefficient formula (one may add a Gaussian remainder to the subset weights). Total degree before the projection is the sum of the tensor degrees. Consequently $\|\mathcal{N}R_+\|_2 \le C_m\|\mathcal{N}\mathcal{T}_{3q}g\|_2$ if $\rho \le 1$. The available strict margin $9q^2 \le a_-/2$ gives

$$\|\mathcal{N}\mathcal{T}_{3q}g\|_2^2 \le C\rho D_-.$$

To see this, write the squared norm as $\sum_n n^2(3q)^n\|g_n\|_2^2$. Use Cauchy–Schwarz against $na_-^{n/2}$ and absorb the remaining $n(3q/\sqrt{a_-})^n$ by its bounded supremum.

Gaussian hypercontractivity [11] gives

$$\|R\|_4 + \|\mathcal{N}R\|_4 + \|\nabla R\|_{L^4(\ell^2)} \le C\|\mathcal{N}R_+\|_2.$$

For the gradient, commute it with the OU operator and use vector-valued conditional Jensen followed by scalar hypercontractivity. Also $\|\nabla R_+\|_2 = \|\mathcal{N}^{1/2}R_+\|_2$. The Gaussian number chain rule $\mathcal{N}\psi(R) = \psi'(R)\mathcal{N}R - \psi''(R)|\nabla R|^2$ and the displayed bounds yield

$$\|\psi(R)\|_2 + \|\mathcal{N}\psi(R)\|_2 \le C_m\rho D_-.$$

Pair with M'' in L^2 and with each L_i'' through $\mathcal{N}^{-1}$. These estimates bound all remaining logarithmic terms by $C_m\rho E$.

The exact entropy second derivative is

$$F''_{q,g}(w) = -\langle L'' + M'',\, L + M + \psi(R)\rangle - \int \frac{(R')^2}{1+R}\, d\gamma_d.$$

The reference Gaussian logarithm drops out because the output is isotropic. The coefficient derivatives and all Gaussian spatial derivatives used here exist in the indicated norms by positive Gaussian smoothing and the displayed degree bounds. In particular the entropy differentiation holds locally in weight by L^2 differentiability and a uniform positive ratio. The last integral is nonnegative, and the first pairing is at least $E - C_m\rho E$. Choose ρ_m with $C_m\rho_m \le 1/2$. The positive signal and a nonzero tangent make (99) strictly positive when $g \ne 0$. This proves Theorem 18.1.

18.4 Several light weights at a fixed interior point

We use the test inputs of Section 17.4. We give the estimates that transport that exact three-copy construction to a fixed point with $m-1$ light weights. On the line let $r \ge 4$ be even, $L \to \infty$ even, and e_j the normalized Gaussian Hermites. The compact even smooth approximants f_L, h_L there satisfy: f_L has no degree below L, and h_L has only the coefficient 1 at r below L. They approach e_L, e_r respectively with exponentially small errors in every fixed Gaussian number Sobolev norm

and bounded C^s norm at most $C_s(L+1)^{C_s}e^{C_1 L}$. These properties come from a cutoff at $C\sqrt{L}$, the Gaussian Hermite L^4 bound, and inversion of the truncated Gram matrix. We use them with the same proof for arbitrary fixed m.

Choose $Q = q/2$, $\delta_0 = cQ^2$, $\epsilon_L = e^{-AL}$ with $A > C_1 + 1$, and $\lambda_L = \lambda(1-\delta_0)^{L-r/2}L^{r/2}$. The latent relative perturbation is the OU image at signal $1/2$ of

$$\epsilon_L f_L + \lambda_L \epsilon_L^2 Q^{L-r/2} h_L.$$

It goes to zero in every fixed bounded relative C^k norm and is strongly log-concave for large L. Its moments below r agree with the Gaussian. For the input after smoothing at q let $M_r = \kappa_r(p_q)$. The construction in Section 17 gives

$$M_r = \sqrt{r!}\,\lambda_L \epsilon_L^2 Q^L > 0,$$
$$B_r(p_{\mathrm{heavy}}) \le -c_r \epsilon_L^2 [Q(1-\delta)]^L L^{r/2}, \qquad 0 \le \delta \le \delta_0.$$

Here $B_r(p) = \int p^{(r)} \log p$ and p_{heavy} is the law with one heavy copy and the remaining inputs Gaussian. These identities and the sign do not depend on the number of light copies. The sign comes from the negative quadratic coefficient $-\frac{1}{2}\langle \mathrm{He}_r, e_L^2\rangle_\gamma$ of B_r and a sufficiently small fixed $\lambda > 0$ for the added linear coefficient at r.

Let $\ell = m - 1$, and write the weights as $(1-\delta, \delta\tau_1, \dots, \delta\tau_\ell)$ with τ a simplex weight. We bound second derivatives along a light tangent v, with $\sum_i v_i = 0$, in τ-coordinates. If $j \ge 4$ is a nonconstant light index, a split derivative of order at most two on $\tau_i^{j/2}$ leaves a nonnegative exponent. For a multinomial term of total degree N, the directional derivatives of orders $s \le 2$ on the product of the split powers cost at most $C_m(N+1)^s|v|^s$. Thus these derivatives are controlled at all split weights.

Use the same decomposition into light coefficient classes as in Section 17. Terms with a light index $\ge L$ admit a reference Gaussian projection with weight $c_1 = 1/(4\ell)$ on each light slot and unchanged weight $Q(1-\delta)$ on the heavy slot. The reference sum is at most $3/4$. In at least one light slot the effective gain $Q\delta\tau_i$ is smaller by a factor giving $(Q\delta/c_1)^{L/2}$; for split derivatives the remaining split powers are still bounded by one. Multinomial Cauchy–Schwarz and the fixed-order Gaussian Sobolev bounds show that all high light terms, including after directional derivatives of order $s \le 2$, have L^2 norm bounded by

$$C_m|v|^s \,\mathrm{poly}(L)\epsilon_L (Q\delta/c_1)^{L/2}.$$

Here and below the fixed polynomial may depend on m, r. For terms where all occupied light slots are at degree r, shifting the heavy coefficient series by kr if k light slots are occupied costs at most a constant times a fixed Gaussian number Sobolev norm of the heavy ratio. That norm is uniformly bounded. The k-slot term and its derivatives of order $s \le 2$ therefore cost at most $C_{m,r}|v|^s(M_r\delta^{r/2})^k$. The singleton r term is kept exactly; all $k \ge 2$ enter the error.

Write $P_r(\tau) = \sum_i \tau_i^{r/2}$. Pair the singleton second derivative with the heavy logarithm using the Hermite raising identity, which sends a degree-r light coefficient to $p_{\mathrm{heavy}}^{(r)}/(\sqrt{r!}\phi)$. The retained term of the entropy Hessian equals

$$-\frac{M_r B_r(p_{\mathrm{heavy}})}{r!}\delta^{r/2}\,\partial_v^2 P_r(\tau).$$

At equal split $\tau_i = 1/\ell$ this is positive, with $\partial_v^2 P_r = k(k-1)\ell^{2-k}|v|^2$ where $k = r/2$. The error from the high contribution paired with the heavy logarithm is bounded by

$$|v|^2 \,\mathrm{poly}(L)\epsilon_L^2 Q^L (\delta/c_1)^{L/2}.$$

Indeed the log has norm at most $\mathrm{poly}(L)\epsilon_L Q^{L/2}$. The high squared norm in the entropy remainder is bounded by the same quantity if $\delta < c_1$. The multiple-low-slot terms and the low squared norm are at most $C_{m,r}|v|^2 M_r^2 \delta^r$. These statements also control replacing the output logarithm by the heavy logarithm in the pairing: both relative densities are bounded above and below by fixed constants, so Cauchy–Schwarz and the L^2 bounds control the difference by the sum of the low and high squared bounds. The entropy derivative square has the same estimate.

Relative to the retained term at equal split, the low error is at most $C_{m,r}\lambda\delta^{r/2}$. The high error, for $0 < \delta \le \delta_0$, is at most

$$\mathrm{poly}(L)\delta_0^{-r/2}\left[\frac{e^{2A}(\delta_0/c_1)^{1/2}}{Q(1-\delta_0)^2}\right]^L.$$

As in Section 17, choose c in $\delta_0 = cQ^2$ so the bracket is below $1/2$. For each fixed positive q the error tends to zero, uniformly on unit light tangents. Take $\delta = \delta_0$. The positive curvature in τ-coordinates gives positive curvature on the light weight subspace by dividing by δ^2. Embedding the scalar law as one coordinate with independent Gaussian coordinates preserves the exact moment matching and adds a weight-independent entropy term. This proves Theorem 18.2.

19 The exact binary Gaussian conditional-product region

19.1 Main result

Let T_b denote the Ornstein–Uhlenbeck operator with multiplier $b^{n/2}$ on Gaussian Hermite degree n, and let $\mathcal{N}$ have multiplier n. All Gaussian norms use the standard Gaussian γ_d. For independent standard Gaussian vectors Z_1, Z_2, G, define

$$M(g_1,g_2)(y) = \mathbb{E}[g_1(Z_1)g_2(Z_2) \mid \sqrt{r}\,Z_1 + \sqrt{s}\,Z_2 + \sqrt{1-r-s}\,G = y],$$

and write $M(g) = M(g,g)$. Order the positive loads as $s \le r$ and put $q = r + s < 1$. Define

$$r_c = (\sqrt{17}-3)/2, \quad s_c(r) = \begin{cases} 6 + 4r - \sqrt{(6+4r)^2 - 4}, & r \le r_c, \\ r(1-r), & r \ge r_c. \end{cases}$$

Theorem 19.1. *For fixed $d \ge 1$, $0 < s \le r$ and $r + s < 1$, there is a finite constant $C_{r,s,d}$ such that every real bounded g satisfies*

$$\|\mathcal{N}M(g)\|_2 \le C_{r,s,d}\|g\|_\infty \|(\mathcal{N}+1)T_r g\|_2$$

if and only if $s \le s_c(r)$. For $s > s_c(r)$, even the bound with $\|M(g)\|_2$ on the left fails in dimension one. Counterexamples can be taken smooth and even with arbitrarily small supremum norm, with strongly log-concave density $\gamma_1(1+g)$, and with any prescribed finite Gaussian moments matched exactly. A witness in dimension one lifts to every fixed dimension by dependence on the first coordinate alone.

For independent X_1, X_2 with density $\gamma_d(1+g)$, define

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

Here g is admissible when it is even and bounded and the density has standard Gaussian mass and covariance.

Corollary 19.2. *If $0 < q < 1$, $t \in [0,1]$, $r = q\max(t, 1-t)$, $s = q\min(t, 1-t)$ and $s \le s_c(r)$, then there are $c, \delta > 0$ such that every admissible g with $\|g\|_\infty \le \delta$ satisfies*

$$F''_{q,g}(t) \le -c\|\mathcal{N}T_r g\|_2^2.$$

At endpoint weights the derivative is one-sided. The constants can be chosen uniformly on compact subsets avoiding the single ordered-load point $(r_c, r_c(1-r_c))$. The earlier dimension-uniform estimate applies at that point itself.

Luo's one-factor oscillator comparison and entropy counterexamples are in [7]. The Evidence Press scalar result covers all weights through $q = 4/7$ [5]. The Luminary Archive gives a fixed-dimensional conditional product estimate near the equal allocation, a dimension-uniform unequal-weight region, and an exact product obstruction when $r \ge (\sqrt{5}-1)/2$ (Sections 11.3, 13 and 16.4). Theorem 19.1 determines the full ordered-load region. For example, Corollary 19.2 applies at $q = 3/5, t = 2801/6000$, outside both the earlier dimension-uniform region and the sufficient comparison with equal input smoothing levels $b_1 = b_2 = q$. The full entropy Hessian includes additional terms, and product failure alone does not fix its sign.

19.2 Allocation and the quadratic criterion

Put $f_{r,s}(u) = (2+r)u^2 + (s-2r-2)u + r + s$ and $u_0 = (2+2r-s)/(2(2+r))$. Use the larger linear load $a = r$ and allocate u of the output to the larger factor. Then the two input smoothing levels are $b_1 = r/u$ and $b_2 = s/(1-u)$, with $r \le u \le 1-s$. Set $h = \sqrt{1-a^2}$, $D = 1+h$, $\lambda = a/D$, and

$$\tau(b) = \frac{h-1+b}{D-b}, \qquad \nu(q) = \frac{1+q-h}{1+q+h}.$$

For $b_i \ge a$, substitution and clearing positive denominators give

$$\tau(b_1)\tau(b_2) \le \lambda \quad \Longleftrightarrow \quad f_{r,s}(u) \le 0.$$

Indeed the difference of the two cleared numerators has the same sign as $f_{r,s}$, with multiplier $a(D-a)$ after also clearing the denominator D in λ. The quadratic minimum on $[r, 1-s]$ is at $\max(r, u_0)$. The upper endpoint does not bind at positive loads when $q < 1$, since $u_0 < 1-s$ follows from $(3+2r)s < 2$, using $s \le \min(r, 1-r)$. At $u = r$, the condition is $s \le r(1-r)$.

The earlier dimension-uniform theorem Section 13 covers $s \le 1/4$ for $r \le 1/2$, or $s \le r(1-r)$ for $r \ge 1/2$. Outside that region, a qualifying minimizer satisfies $r < u_0 < 1-s$, and $b_1, b_2 > a$. For $r \le 1/2$, use $u_0 \ge 1/2 > 1 - s/r$, as $s > 1/4$. For $r \ge 1/2$, use $1 - s/r < r$. In both cases $b_1 > a$ follows from $u_0 < 1$. The discriminant condition there is $4(2+r)(r+s) \le (2+2r-s)^2$. In the example in Section 1 the loads are $r = 3199/10000$, $s = 2801/10000$. $f_{r,s}(51/100) = -4101/10^8$, and $u = 51/100$ meets all strict level constraints. At the half allocation $f_{r,s}(1/2) = 1/8000 > 0$. Equal smoothing levels instead give $u = r/q = 3199/6000$, where $f_{r,s}(u) > 0$ as well.

19.3 A collective oscillator estimate

We use the conditional-product lemma in Section 11.3, from its collective shell estimate through the end of that lemma. Its proof assumes a neighborhood of equal weights. The following argument replaces the only Gaussian decay step using that neighborhood assumption. It works for any fixed positive weights w_i, $q > a$, and $a \le b_i < 1$ with $\sum_i q w_i/b_i = 1$. For the binary theorem $w = (r/q, s/q)$. The constants below are uniform on compact admissible level and positive-weight neighborhoods where the strict matrix gaps hold.

Write $r_0 = \lambda^2$, $c_i = \sqrt{w_i}$, and $H = \operatorname{diag}(\tau(b_i))$. In the prior shell proof the normalized replica matrices are

$$P = H^{-1/2}(r_0 I + (\nu(q) - r_0)cc^\top)H^{-1/2}, \qquad Q = H^{-1/2}(-r_0 I + (\tau(q) + r_0)cc^\top)H^{-1/2}.$$

The identities $\tau(b) + r_0 = 2hb/[D(D-b)]$ and $\sum_i qw_i/b_i = 1$ give $I - Q \succeq 0$ with exactly one null direction e. Also $Q > -I$ follows from $\tau(b_i) \geq \lambda > r_0$. The matrix P is positive. To prove $P < I$ use

$$\tau(b) - r_0 = \frac{2(b - a^2)}{D(D-b)}, \qquad \nu(q) - r_0 = \frac{2hq}{D(D+q)}.$$

The rank-one test for $I - P > 0$, with $u_i = qw_i/b_i$, therefore reduces to

$$\frac{h}{D+q} \sum_i u_i \frac{b_i(D-b_i)}{b_i - a^2} < 1.$$

The function $b(D-b)/(b-a^2)$ decreases for $b \geq a$, because the derivative numerator is $-(b-a^2)^2 - a^2(D-a^2)$. Its value at a is $(D-a)/(1-a)$. Since $h(D-a)/(1-a) = D + a$, the displayed left-hand side is at most $(D+a)/(D+q) < 1$ for $q > a$.

The physical Gaussian representation in the prior proof now has every positive covariance eigenvalue strictly below $1/B_a$, where $B_a = 1/h$. In fact the sum and difference covariances are $(I+P)/(2B_a)$ and $(I-Q)/(2B_a)$. Its support is the subspace on which the collective coordinate $s_0 = e^\top U = e^\top V$ is shared. The transverse replicas are conditionally independent given s_0, since the cross covariance is rank one and s_0 is shared with positive variance. Dividing out the oscillator ground states multiplies this Gaussian by $\exp(B_a(|U|^2 + |V|^2)/2)$. The covariance gaps imply a strict Gaussian bound on that whole supporting subspace, even when the conditional transverse mean is nonzero.

More explicitly let p be the density of s_0, p_{s_0} the transverse conditional density and $\phi_{s_0}(z) = e^{B_a|z|^2/2} p_{s_0}(z)$. The preceding bound, evaluated on $U = (s_0, z)$, $V = (s_0, z)$ and integrated in z, gives

$$p(s_0) e^{B_a|s_0|^2} \|\phi_{s_0}\|_2^2 \leq C e^{-\eta|s_0|^2}.$$

The conditional Gaussian has fixed strictly positive precision after multiplication by $e^{B_a|z|^2/2}$, by the same joint bound on the diagonal. The Bargmann series of ϕ_{s_0} is a scalar depending on s_0 times $\exp(\zeta^\top A\zeta/2 + \ell_{s_0}^\top \zeta)$, where $\|A\| < 1$ and the linear coefficient depends linearly on s_0. Replacing ζ by $\zeta/\sqrt{\sigma}$ applies the inverse oscillator attenuation $\sigma^{-\mathcal{H}_\perp/2}$. For $\sigma < 1$ close enough to one the Gaussian Bargmann norm still converges. Completing the square shows that the resulting quadratic coefficient in s_0 is continuous in σ. The strict decay in the last display thus persists with perhaps smaller $\eta > 0$. This gives the bound of the prior proof with $e^{-\eta'|s_0|^2}\sigma^{\mathcal{H}_\perp/2}$ for its normalized Gram.

All subsequent estimates in the prior proof apply with these constants. In detail, separation by transverse degree and the localized collective shell estimate give the $(N+1)^{-1/4}$ gain. The spherical tensor estimate is uniform in e and its multinomial and Gamma bounds require no symmetry in the weights. Multiplying all b_i and q by a common factor preserves the allocation, and the matrix gaps persist locally under that change. The upward change by $\exp(c_0/(N+1))$ and integrated downward change therefore give the output number and inverse number bounds. Their uniform tensor estimates may be applied before integration. The input weight derivatives cost two input number operators, and the inverse output number leaves the identical factor $C(N+1)$. Finally, the shell summation uses only $\prod_i \tau(b_i) \leq \lambda$, levels bounded away from zero, and $\max_i \tau(b_i) < 1$. The cutoff and Gaussian L^2 approximation work as before. Thus for two factors at the stated allocations and $Y = \|g\|_\infty$,

$$\|\mathcal{N} M\|_2 + \|\mathcal{N}^{-1} M''\|_2 \leq CY \, \|(\mathcal{N}+1) T_a g\|_2,$$

where M is the mixed density term. Weight derivatives are taken at the fixed physical weights; one does not differentiate an optimizing allocation.

19.4 Sharpness of the product estimate

19.4.1 Choosing the replica allocation

Assume $f_{r,s}(\max(r,u_0)) > 0$. The earlier dimension-uniform allocation region is contained in the quadratic region: for $r \le 1/2, s \le 1/4$ use $f_{r,s}(1/2) = -1/2 + r/4 + 3s/2 \le 0$, and for $r \ge 1/2, s \le r(1-r)$ use $f_{r,s}(r) \le 0$. Thus, if $r \le 1/2$ then $s > 1/4$, while if $r > 1/2$ then $s > r(1-r)$.

Minimize $J(u) = \log(\tau(b_1)\tau(b_2))$ on $[\max(1/2, r), 1-s]$. On this interval $a \le b_i \le 1$, since the preceding alternatives give $b_2 > a$ at the left endpoint, b_2 increases, and $b_1 = r/u \ge a$. Every allocation has $\tau(b_1)\tau(b_2) > \lambda$. Let

$$G(b) = \frac{b^2\tau'(b)}{\tau(b)} = \frac{2hb^2}{-a^2 + 2b - b^2}.$$

Then $J'(u) = -G(b_1)/r + G(b_2)/s$. Both $G(b)$ and $G(b)/b$ are nondecreasing for $b \ge a$, with G strictly increasing. At $u = 1-s$, $b_2 = 1 > b_1$, so $J'(u) > 0$. At $u = 1/2$ when $r \le 1/2$, the second monotonicity gives $J'(u) \le 0$. Consequently a minimizer either has $J' = 0$, or occurs at $u = r$, where $b_1 = 1$ and $J' \ge 0$. This description includes equal ordered loads.

For the normalized Gram, the collective null vector of $I - Q$ can be taken positive. Its squared coordinates satisfy

$$e_i^2 \;\propto\; \frac{w_i\tau(b_i)}{(\tau(b_i) + r_0)^2} \;\propto\; \frac{qw_i}{G(b_i)}.$$

The second identity follows from the formula for $\tau + r_0$ and the derivative of τ. Therefore $J' = 0$ gives $e_1 = e_2$. In the only other case $b_1 = 1$, $b_2 < 1$ and $e_1 \ge e_2 > 0$. The replica Gram and its strict Gaussian bounds apply at $b_1 = 1$ as well: the rank-one inequalities above allow $b_i = 1$, and the determinant in the generating calculation depends only on q, w, a. No summation at a possibly unit $\tau(b_i)$ is used for necessity.

19.4.2 Oscillator tensors realizing the boundary

In dimension one let H_n be the $L^2(dx)$-normalized oscillator function of frequency B_a. For even n choose its sign so $H_n(0) > 0$. Put $k_n = \sqrt{(2n+1)B_a}$. The oscillator equation with even initial data gives

$$\frac{H_n(x)}{H_n(0)} = \cos(k_n x) + O\!\left(\frac{R^3}{k_n}\exp(CR^3/k_n)\right) \qquad (|x| \le R),$$

by the sine-kernel Duhamel formula for $H_n'' + k_n^2 H_n = B_a^2 x^2 H_n$. Stirling gives $H_n(0) \asymp (n+1)^{-1/4}$. The energy identity gives $\|H_n'\|_2 \le k_n$, and the one-dimensional Sobolev inequality gives $\|H_n\|_\infty \le C(n+1)^{1/4}$. The Duhamel error is therefore uniformly $o(1)$ for $R = C_1\sqrt{\log(n+2)}$, and the ratio $H_n(x)/H_n(0)$ has a polynomial global bound.

Choose even $n_2 \to \infty$ and choose even n_1 such that $k_{n_1}e_2 - k_{n_2}e_1 = O(n_2^{-1/2})$, by rounding the matching oscillator energy to the nearest even index. When $e_1 = e_2$ take $n_1 = n_2$. At $b_1 = 1$ with $e_1 > e_2$ one has $n_1/n_2 \to (e_1/e_2)^2 > 1$. Denote the bilinear conditional product by $M(h_1, h_2)$. For the tensors $H_{n_1} \otimes H_{n_2}$ the exact normalized Gram is the square-integral of the conditional transverse pairing with ϕ_{s_0}, up to the same fixed positive prefactor. In the rotated variables

$x_1 = e_1 s_0 - e_2 z,\ x_2 = e_2 s_0 + e_1 z$. On any fixed compact s_0-interval, divide the transverse integral by $H_{n_1}(0) H_{n_2}(0)$. The Gaussian bound on ϕ_{s_0} and the polynomial global oscillator bounds allow a cutoff at a sufficiently large fixed multiple of $\sqrt{\log(n_2+2)}$ with uniform $o(1)$ error. The Duhamel estimate inside the cutoff then gives the integral of a product of two cosines, plus a uniform $o(1)$.

The sum phase has transverse frequency $k_{n_2} e_1 - k_{n_1} e_2 = O(n_2^{-1/2})$. The difference phase has transverse frequency of order $\sqrt{n_2}$. The Gaussian transform for ϕ_{s_0} therefore yields

$$\frac{1}{2} A(s_0) \cos((k_{n_1} e_1 + k_{n_2} e_2) s_0) + o(1), \qquad A(s_0) = \int \phi_{s_0}(z)\, dz > 0.$$

On a fixed positive-length interval, the weighted squared norm of this expression is bounded below by a positive constant, by the oscillatory average of the cosine squared. Restoring the input attenuation yields

$$\|M(H_{n_1}, H_{n_2})\|_2 \geq c\, (n_1+1)^{-1/4} (n_2+1)^{-1/4} \tau(b_1)^{n_1/2} \tau(b_2)^{n_2/2}.$$

For the reverse tensor the normalized Gram is a bounded $L^2(dx)$ form by the joint Gaussian margin and transverse Cauchy–Schwarz. This gives the upper bound $C\tau(b_1)^{n_2/2}\tau(b_2)^{n_1/2}$ for its output norm.

For the energy on the right, the single-factor oscillator Gram is diagonal at a, with value $C\lambda^n$ on shell n. At $b = a\exp(c/(n+1)) < 1$ its Gaussian multiplication bound gives $\|T_b H_n\|_2 \leq C\tau(b)^{n/2} \leq C'\lambda^{n/2}$. The final smoothing from b to a then pays a factor $C(n+1)$ for the number derivative. Thus $\|(\mathcal{N}+1) T_a H_n\|_2 \leq C(n+1)\lambda^{n/2}$.

The equal-index case gives a single-input witness immediately, because $\tau(b_1)\tau(b_2) > \lambda$. In the unequal-index case, $\tau(b_1) = 1$ and $\tau(b_2) > \lambda$. The reverse tensor is exponentially smaller than the displayed lower bound. If the asserted comparison with $\|M\|_2$ were true for every input, apply it to $H_{n_1} + H_{n_2}$ and $H_{n_1} - H_{n_2}$ and subtract their quadratic products. The symmetric mixed product on the left has the displayed lower bound up to constants. The right is at most a polynomial times $\lambda^{n_2/2}$, which is exponentially smaller.

19.4.3 Exact moments

For each fixed even moment degree $2j$, the inner product of H_n with x^{2j} under the standard Gaussian is bounded by a polynomial in $n+1$ times λ^n. Indeed integration of the oscillator generating function yields a fixed polynomial in its variable times $\exp(\lambda^2 z^2/2)$. Taking its n-th coefficient and multiplying by $\sqrt{n!}$ gives the bound. Subtract a linear combination of $x^{2i} e^{-x^2}$ up to the required moment order, using the invertible weighted polynomial Gram matrix, to annul all prescribed even moments exactly. Odd moments already vanish. Every correction coefficient is exponentially smaller than the relevant lower bound, with rate at most a polynomial times $\lambda^{\min(n_1,n_2)}$. Conditional contraction and the polynomial bounds for the input sup norms show that the correction does not change the product obstruction. The smoothing energy of the corrections has this same smaller rate. Apply polarization to the individually moment-corrected modes. Since $\mathbb{E} M(g,g) = (\mathbb{E} g)^2 = 0$ for every corrected input, the Gaussian number norm is at least the output L^2 norm. The corrections and oscillators each have bounded first two derivatives for every fixed index, so a sufficiently small scalar multiple gives a positive strongly log-concave Gaussian-relative density. The divergent ratio is unchanged by scaling.

This sharp conditional comparison now covers the entire intermediate binary-load range. Alignment of two unequal oscillator degrees is needed when the minimizing allocation is pinned at one unsmoothed input.

19.5 Closed forms for the load and weight regions

Set

$$r_c = (\sqrt{17} - 3)/2, \qquad q_c = 2r_c - r_c^2 = (5\sqrt{17} - 19)/2.$$

Among the physically admissible loads $0 < s \le r,\ r+s < 1$, the exact conditional-product condition is

$$s \le \begin{cases} 6 + 4r - \sqrt{(6+4r)^2 - 4}, & 0 < r \le r_c, \\ r(1-r), & r_c \le r < 1. \end{cases}$$

Indeed the unconstrained quadratic minimum has nonpositive sign exactly when $s^2-(12+8r)s+4 \ge 0$, by expansion. On the admissible range this is equivalent to s below the smaller root shown. The minimum binds at $u = r$ when $f'_{r,s}(r) \ge 0$. At $s = r(1-r)$ this derivative is $r^2 + 3r - 2$. If $r \ge r_c$, any larger s makes the derivative and the minimum positive. If $r < r_c$, the derivative is negative at $s = r(1-r)$. If it first reaches zero within the feasible range, the minimum at that later s is already positive (it equals $f_{r,s}(r)$). Thus the unconstrained sign boundary, whenever in the feasible range, occurs before the binding point. If the binding point is already outside the feasible range, the unconstrained test directly applies where needed.

For each fixed $q \in (0,1)$ this is the exact half-width measured from the nearer endpoint:

$$W_{\mathrm{prod}}(q) = \begin{cases} 1/2, & q \le 4/7, \\ 1 - \dfrac{5q - 6 + 4\sqrt{q(q+3)}}{9q}, & 4/7 \le q \le q_c, \\ \dfrac{\sqrt{1-q}}{1+\sqrt{1-q}}, & q_c \le q < 1. \end{cases}$$

Theorem 19.1 applies at each weight $0 < t < 1$ with $\min(t, 1-t) \le W_{\mathrm{prod}}(q)$, with a number-norm constant allowed to depend on the weight and fixed dimension. Above that width no constant exists even in dimension one. The endpoint weights give the zero mixed term for centered g. For the intermediate piece, substitute $s = q - r$ in the smaller-root equation and take the physical solution $r = (5q - 6 + 4\sqrt{q(q+3)})/9$. Along the first boundary $r + s_c(r)$ increases from $4/7$ at $r = 2/7$ to q_c; this follows from $s_c'(r) = -8s_c/\sqrt{(12+8r)^2 - 16} > -1$. Along the second boundary $r + r(1-r)$ strictly increases to one. Corollary 19.2 applies throughout this interval.

19.6 Entropy curvature

Let $L = T_{qt}g + T_{q(1-t)}g$. The admissible Hermite degrees start at four. Thus $\langle L'', L\rangle \ge c\|\mathcal{N}T_r g\|_2^2$, $\|(\mathcal{N}+1)T_r g\|_2 \le C\|\mathcal{N}T_r g\|_2$, and $\|\mathcal{N}^{-1}L''\|_2 + \|\mathcal{N}L\|_2 \le C\|\mathcal{N}T_r g\|_2$. These follow from the degree multiplier $q^k(t^k + (1-t)^k)$ for degree $2k$ and its second derivative. The Gaussian number-domain chain rule in Section 13 bounds $\|\mathcal{N}[\log(1+L+M) - (L+M)]\|_2$ by $CY\|\mathcal{N}(L+M)\|_2$. Pair the second derivative through $\mathcal{N}^{-1}$ in the entropy Hessian. Its last term is $-\int ((L+M)')^2/(1+L+M)\, d\gamma_d$. Every error to the coercive term has absolute value at most $CY\|\mathcal{N}T_r g\|_2^2$. This proves the strict conclusion when $g \ne 0$.

The Archive's binomial projection estimate Section 13 proves the product bound and its derivative estimate for bounded inputs in the dimension-uniform region. Its entropy theorem also handles the endpoint weights. The new part has uniform constants on compact subsets where both $b_i < 1$. For limits from the new part to the original boundary, the minimizing u_0 can reach r only with $f(r) = 0$, at $(r_c, r_c(1-r_c))$. Away from that point $u_0 < 1 - s$ and $u_0 < 1$. If $r > 1/2$ at such a limit then $1 - s/r = r < u_0$. If $r \le 1/2$, the limit has $s = 1/4$, and $u_0 \ge 1/2 \ge 1 - s/r$ with at least one strict inequality. Thus the limiting levels remain admissible with both $b_i < 1$. The strict

replica gaps persist there and compactness gives the claimed uniformity. At the seam point the old theorem applies. The argument allows constants to grow along approaches from the new part.

20 Log-concave laws on expanding trimmed simplices

20.1 Statement

Let X be isotropic and log-concave in $\mathbb{R}^d$. Let p_s denote the density of $\sqrt{s}X + \sqrt{1-s}\,G$. With independent copies X_i, put

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

Here G is standard normal and independent, h is differential entropy, and w belongs to the probability simplex. Write $w_- = \min_i w_i$, $W = \max_i w_i$, and ϕ for the standard Gaussian density. Use the Hellinger distance $\mathsf{H}(p,r) = \|\sqrt{p} - \sqrt{r}\|_{L^2(dx)}$.

Theorem 20.1. *Fix $m \geq 3, d \geq 1$ and $0 < \alpha < 2$. There are $q_0, c > 0$ depending only on these parameters such that, for $0 < q \leq q_0$ and $w_- \geq q^\alpha$,*

$$\partial_v^2 F_{q,X}(w) \leq -c\sum_i \frac{v_i^2}{w_i^2}\mathsf{H}(p_{q\sqrt{Ww_i}},\phi)^2 \qquad \left(\sum_i v_i = 0\right).$$

For non-Gaussian X and nonzero v this is strict.

Consequently every fixed trimmed simplex admits a common positive signal interval for the entire isotropic log-concave class. The trim may also decrease with q at the rate in the theorem. For centrally symmetric inputs in a uniform Gaussian-relative neighborhood, the sharp trim order is q^2 (Section 18); its counterexamples give the same obstruction to extending the present whole-class theorem too close to the faces. In particular any power larger than two fails for the whole class. Section 6 proves a bounded-relative Hessian estimate at the exact barycentre and Section 7 gives moment-matched counterexamples approaching a face. Evidence Press [5, Section 9.3] asks for a Gaussian-relative neighborhood on a fixed trimmed region. The two-copy theorem of Section 15 covers all weights for centrally symmetric inputs in fixed dimension. Its Fisher interpolation estimate applies without symmetry and is used here.

20.2 Fisher control at the geometric-mean signal

Write $I(p) = \int |\nabla \log(p/\phi)|^2 p$. The fixed-dimensional Fisher estimates of Section 15.2 give $I(p_t) \leq C_d t^2$ uniformly on the isotropic log-concave class for small t. They also give convexity of $\log(I(p_t)/t)$ as a function of $\Theta(u) = u + O(e^{-2u})$, $t = e^{-2u}$, on a uniform small-signal interval. The Fisher bound follows from the relative Hessian bound and Poincaré, since the relative score has mean zero. We use these facts at three signals.

For each i set $s_i = q\sqrt{Ww_i}$, $D_i = \mathsf{H}(p_{s_i},\phi)$. If any $D_i = 0$, Gaussian deconvolution gives Gaussian X, so assume all $D_i > 0$. Let $M > 1$ be a fixed large constant and put $I = I(p_{Mq})$. For every $\beta < 1/(1+\alpha/2)$, after reducing the common signal range,

$$I \leq C D_i^{2\beta} \quad \text{for every } i. \tag{101}$$

To prove this, write $s = s_i$, $D = D_i$, and use $s_\ell = 3s/4$ and a fixed small upper signal s_h within the convexity range. The lower-signal coefficient in interpolating Mq between s_ℓ, s_h is at least $(1+\alpha/2)^{-1} - o(1)$, uniformly. Indeed

$$\log(1/s_\ell) \le (1+\alpha/2)\log(1/q) + O_m(1), \qquad \log(1/(Mq)) = \log(1/q) + O_M(1),$$

and the modified time differs from ordinary logarithmic time by a bounded additive error. The lower-signal coefficient is $(\Theta(u_{Mq}) - \Theta(u_{s_h}))/(\Theta(u_{s_\ell}) - \Theta(u_{s_h}))$. The signal prefactor in the resulting interpolation of I is bounded uniformly: the interpolated ordinary times differ by at most a bounded error from u_{Mq}, so $Mq/(s_\ell^\theta s_h^{1-\theta})$ is bounded. Also $I(p_{s_h})$ is bounded from above.

For every $0 < \zeta < 1$ we have $D_{\mathrm{KL}}(p_s \| \phi) \le C_\zeta D^{2\zeta}$ uniformly for small s. The centered entropy integrand is bounded by $C_d(1+|x|^2)(\sqrt{p_s} - \sqrt{\phi})^2$ because p_s is uniformly bounded by Gaussian smoothing. Hölder and uniform moments absorb the polynomial weight at the cost of ζ. OU entropy dissipation over the fixed time from s to $3s/4$ gives $I(p_{s_\ell}) \le C D_{\mathrm{KL}}(p_s \| \phi)$. Choose ζ close enough to one so that $\zeta\theta > \beta$ uniformly. Interpolation proves (101), using boundedness of D. For each fixed $b_0 < 1$, the same interpolation gives

$$I(p_{Mqw_i}) + I(p_{2s_i}) \le C D_i^{2b_0}. \tag{102}$$

Indeed $Mqw_i \le Ms_i$ and Fisher information is monotone in the signal. For each fixed multiple of s_i the low-level coefficient tends to one uniformly as $s_i \downarrow 0$, by the same time calculation. Also $D_i \le C s_i$ and $I \le C(Mq)^2$ by the relative Fisher estimate and Gaussian logarithmic Sobolev.

Choose $1/2 < \beta < \beta_0 < 1/(1+\alpha/2)$ and $b < b_0 < 1$, with b close enough to one that $b + 2\beta > 2$. Then choose $a < 1$ close enough to one that $e = 2a\beta_0 > 1$ and $b + e > 2$. We apply (101) with β_0. Choose $\gamma < 1$ sufficiently close to one and M sufficiently large that, with $\rho = 1/M$,

$$\gamma/(1+2\rho) > a, \qquad (3/2)\gamma > a, \qquad \rho < 1/(8m).$$

These strict choices will absorb polynomial costs from Gaussian kernel derivatives.

20.3 A bounded Gaussian-relative approximation

Let $\mu = p_{Mq}$, and take $R = (2\gamma \log(1/I))^{1/2}$. We construct a bounded relative perturbation g° by cutting $\mu/\phi - 1$ to $|x| \le R$ and correcting its mass, mean and covariance. Uniformly over the input class,

$$\|g^\circ\|_\infty = o(1), \qquad \|g^\circ\|_{L^2(\phi)} \le C\sqrt{I}. \tag{103}$$

We detail the truncation, since μ/ϕ need not be bounded.

The relative log-density gradient is at most $C(1+|x|)$, by log-concavity, the posterior Hessian cap, and the bounded score at zero. For the last bound, the conditional mean formula at zero and Jensen give a uniform lower bound on the posterior normalizer and a uniform first moment in the numerator. A log-density ratio of absolute size $b \le 1$ at a point in the ball of radius R persists at size at least $b/2$ on a ball of radius $cb/(1+R)$. On this smaller ball the Gaussian density is bounded below by $c_d e^{-R^2/2}$, with a constant independent of its center. The Hellinger squared cost is at least $c\, b^{d+2}(1+R)^{-d} e^{-R^2/2}$. The same test with $b = 1$ rules out larger values once q is small. Since $\mathsf{H}(\mu,\phi)^2 \le CI$ and $e^{-R^2/2} = I^\gamma$, the log-density ratio tends uniformly to zero throughout $|x| \le R$. The cutoff perturbation has $L^2(\phi)$ norm at most $C\sqrt{I}$ by the same Hellinger bound and its now bounded ratio.

On each ray use the two points at radii $R - 1/R$ and R. The log-density ratio at both points is $o(1)$, so the radial logarithmic slope beyond R is at most $-(1-o(1))R$ by log-concavity. Uniformly over directions θ, for $u \geq R$ we have

$$\mu(u\theta) + \phi(u\theta) \leq C_d I^\gamma e^{-(1-o(1))R(u-R)}. \tag{104}$$

Thus the cutoff mass, mean and covariance errors are $O(\mathrm{poly}(R)I^\gamma)$. Correct them by adding a linear combination of fixed smooth compactly supported functions. Take a nonnegative cutoff positive on a ball times the polynomials $1, x_i, x_i x_j$ for $i \leq j$. Their Gram matrix under ϕ is positive definite by polynomial independence on that ball, so the coefficients have the same small bound. This enforces the exact mass, mean and covariance. Since $\gamma > 1/2$, the correction also proves (103).

20.4 Weighted approximation of the densities

It is convenient to use nearby off-simplex weights with Gaussian variance chosen to keep covariance I_d. All estimates below are evaluated at the original w. For each subset J, let p_J denote the output with latent copies in J and Gaussian copies elsewhere, so $p_\emptyset = \phi$ and p_{all} is the true output. At the common pre-signal Mq, the effective squared gains are ρw_i and the Gaussian spare variance is $1 - \rho \sum_i w_i$. Put

$$\mathcal{R} = 2^{-m} \sum_J p_J, \qquad L_i = \mathcal{T}_{\rho w_i} g^\circ, \qquad p_{0,J} = \phi \left(1 + \sum_{i \in J} L_i\right),$$

where $\mathcal{T}_b = T_{\sqrt{b}}$ has multiplier $b^{n/2}$ on degree n. The mixture $\mathcal{R}$ is obtained by choosing the input in each slot from $\bar{\mu} = (\mu + \phi)/2$. All model ratios $p_{0,J}/\phi$ are uniformly close to one.

For a signed density f, write

$$\|f\|_{K,\mathcal{R}} = \left(\int (1 + |x|)^K f(x)^2 / \mathcal{R}(x)\, dx\right)^{1/2}.$$

We will use fixed sufficiently large K; all estimates below apply to every fixed K. Normalize weight derivatives as $w_i \partial_i$ or $w_i w_j \partial_{ij}$, with the prefactors evaluated at w. In a Gaussian output kernel the normalized derivatives through order two contribute only a uniformly bounded-coefficient polynomial in the latent norms and the Gaussian innovation. For instance $w_i \partial_i \sqrt{\rho w_i} = \sqrt{\rho w_i}/2$ and $w_i^2 \partial_i^2 \sqrt{\rho w_i} = -\sqrt{\rho w_i}/4$. Variance derivatives are uniformly bounded too, with all spare variances bounded away from zero.

Let ℓ_i be the signed density difference $p_{\{i\}} - \phi$, and $t_i = \ell_i - \phi L_i$. We claim

$$\|t_i\|_{K,\mathcal{R}} + \|w_i^2 \partial_i^2 t_i\|_{K,\mathcal{R}} \leq C_K I^a. \tag{105}$$

We also use this zero-order estimate at any synthetic singleton level $0 < b_1 \leq 2\rho$ in the two-component output mixture, which is half the sum of ϕ and the true singleton density at that level.

Indeed each singleton has channel kernel of variance $1 - b_1$ and center $\sqrt{b_1} y$. The compact correction from (103) costs $O(\mathrm{poly}(R)I^\gamma)$, since the denominator dominates a fixed multiple of ϕ. On $R \leq |y| \leq 2R$, completing the square shows that the kernel, or its normalized second weight derivative in a true slot, has norm in $L^2((1 + |x|)^K \phi^{-1} dx)$ bounded by

$$C_K (1 + |y|)^C \exp\left(\frac{b_1 |y|^2}{2(1 + b_1)}\right).$$

Derivatives add only polynomial factors, uniformly in $b_1 \le 2\rho$. Integrate this norm against $(\mu + \phi)1_{\{R\le|y|\le 2R\}}$ using (104). The exponential in the radius has maximum at R, because $2b_1/(1+b_1) < 1-o(1)$. The result is $O(\mathrm{poly}(R)I^{\gamma/(1+2\rho)})$. For $|y| > 2R$, represent the signed tail in the appropriate mixture conditional expectation and use conditional Cauchy–Schwarz. The latent ratio is bounded by two and supported on that tail, so the resulting norm is bounded by

$$C_K \left(\int_{|y|>2R} (1+|y|)^C \bar{\mu}(y)\, dy \right)^{1/2} \le C_K \,\mathrm{poly}(R) I^{(3/2-o(1))\gamma}.$$

For the m-slot mixture one may represent the original singleton derivative with all other slots Gaussian by the bounded Gaussian ratios $\phi/\bar{\mu}$. Integration in the independent slots and the Gaussian innovation absorbs the other fixed polynomial moments. For the synthetic mixture the same estimate uses one slot. Since $\gamma/(1+2\rho) > a$ by choosing ρ still smaller if necessary, all polynomial factors are absorbed and (105) follows.

The same conditional argument handles signed products in the exact expansion of p_J into subsets of $\mu - \phi$ slots. For a signed slot with original density μ_1 and mixture $r_1 = (\mu_1 + \phi)/2$, one has

$$\int (\mu_1 - \phi)^2/r_1 \le CI(\mu_1), \qquad |\mu_1 - \phi|/r_1 \le 2.$$

Hölder with uniform polynomial moments makes the first inequality hold with any fixed polynomial weight at the cost of replacing I by I^ζ, for any fixed $\zeta < 1$. After differentiating a kernel up to order two, conditional Cauchy–Schwarz bounds the squared reciprocal norm by moments of the squares of the inserted ratios times this kernel polynomial. Bound the polynomial by a product of fixed polynomial weights and then use independence. Each signed slot therefore contributes a factor $I(\mu_1)^{\zeta/2}$ in the norm. With all pre-signals equal to Mq, every product of at least two signed slots costs $O(I^\zeta)$. Together with (105) and $\zeta > a$ this yields, for every subset J,

$$\|p_J - p_{0,J}\|_{K,\mathcal{R}} \le C_K I^a. \tag{106}$$

For the derivatives of the nonlinear terms we use a different representation of the same densities at the evaluation point. Write $p_{\mathrm{all}} = \phi + \sum_i \ell_i + \mathcal{B}$, where $\mathcal{B}$ is the sum over signed subsets H with $|H| \ge 2$. A term indexed by H depends only on the gains of its occupied slots: the Gaussian inputs in the others can be integrated out. Thus ∂_{ij} acts only if both $i, j \in H$, with the case $i = j$ included. Fix an evaluation weight w. In slot ℓ pre-smooth at $u_\ell = Mqw_\ell$, held fixed when taking weight derivatives. Its squared gain at a nearby w' is qw'_ℓ/u_ℓ, equal to ρ at the evaluation point. The spare variance after adding the active slots is positive near this point since $m\rho < 1/8$. To use the mixture for a differentiated subset, keep the gains of its inactive Gaussian slots fixed at ρ and reduce the spare variance accordingly. At the evaluation point the distribution of the full reference mixture is still $\mathcal{R}$: every choice of true or Gaussian pre-inputs yields exactly the same p_J after combining all independent Gaussian variances. Normalized derivatives of the individual gains and variance are bounded in this representation at w.

In each differentiated subset H, slot i is signed, and at least one other slot is signed. Use (102) for that i and (101) with β_0 for the others, since all their pre-signals are at most Mq. Choose the Hölder exponent $\zeta < 1$ close enough to one to preserve $b < b_0$ and $\beta < \beta_0$ in the norm bound. The conditional argument gives

$$\|w_i w_j \partial_{ij} \mathcal{B}\|_{K,\mathcal{R}} \le C_K D_i^{b+\beta}. \tag{107}$$

The same bound holds with i, j interchanged. This representation estimates the raw terms; the bounded linear model still uses the same g° in every slot.

20.5 Comparing the entropy pairings

Write $p = p_{\text{all}}$, $p_0 = p_{0,\text{all}}$, $L = \sum_i L_i$, and $\ell = \sum_i \ell_i$. All true output densities have a Gaussian lower envelope and uniform upper bound, by restricting isotropic pre-inputs to a ball of fixed positive mass and using the spare Gaussian variance. Thus the logarithms of their ratios to ϕ are bounded by $C(1+|x|)^2$. The same bound applies to $\log(p/p_0)$.

In $L^2(\mathcal{R}\,dx)$ we have

$$\|\log(p/\phi)\|_{L^2(\mathcal{R})} \le C\sqrt{I}, \tag{108}$$

$$\|\log(p/\phi) - L\|_{L^2(\mathcal{R})} \le CI^a. \tag{109}$$

For the latter, first compare p and p_0. Where $\mathcal{R} \le C_0(p+\phi)$, the elementary logarithm difference bound is

$$|\log(p/p_0)| \le C(1+|x|)^2|p-p_0|/\mathcal{R}.$$

Indeed p_0/ϕ is uniformly bounded above and below, and the logarithm is Lipschitz when p/p_0 is between $1/2$ and 2. On its complement $|p-p_0|/(p+p_0)$ has a fixed positive lower bound. Where $\mathcal{R} > C_0(p+\phi)$, some subset density $p_J \ge \mathcal{R}$, and, for C_0 large, $|p_J - p_{0,J}| \ge \mathcal{R}/2$ since $p_{0,J} \le C\phi$. Thus the squared integral of the polynomial log bound on this bad set is charged to the sum of weighted subset errors in (106). This proves $\|\log(p/p_0)\|_{L^2(\mathcal{R})} \le CI^a$.

On $\mathcal{R} \le C_0\phi$, the norms of $\log(p_0/\phi)$ and of $\log(p_0/\phi) - L$ are at most $C\|L\|_2$ and $C\|L\|_4^2$, respectively, with Gaussian norms. On the complement charge the same subset errors, since these functions are bounded there too. Gaussian hypercontractivity [11] with $\rho < 1/8$ gives $\|L\|_2 + \|L\|_4 \le C\|g^\circ\|_2 \le C\sqrt{I}$. As $a \in (1/2, 1)$, this proves (108)–(109), and $\|\log(1+L) - L\|_2 \le CI$ in Gaussian norm.

Next consider any synthetic level b_1 with $0 < b_1 \le 2\rho$. Let $p_{s'}$ be the true singleton at $s' = Mqb_1$ and $\Delta_{b_1} = p_{s'} - \phi(1 + \mathcal{T}_{b_1} g^\circ)$. In its two-component mixture $\mathcal{R}_{b_1} = (p_{s'} + \phi)/2$, the estimates of the previous section give

$$\|\Delta_{b_1}\|_{K,\mathcal{R}_{b_1}} \le C_K I^a.$$

The Hellinger distance between the two densities is therefore at most CI^a, because the model ratio is bounded above and below. For $b_i = \rho\sqrt{W w_i}$, we have $s' = s_i$, so the triangle inequality gives

$$\|\mathcal{T}_{b_i} g^\circ\|_2 \ge cD_i. \tag{110}$$

Here $I^a \le CD_i^e = o(D_i)$. For upper bounds, the same comparison and (102) imply

$$\|\mathcal{T}_{b_1} g^\circ\|_2 \le CD_i^b \qquad (b_1 \le 2b_i). \tag{111}$$

Indeed $\mathsf{H}(p_{s'}, \phi) \le \mathsf{H}(p_{2s_i}, \phi) \le C\sqrt{I(p_{2s_i})}$ by contraction and logarithmic Sobolev, and $I^a = O(D_i^e)$ is smaller. The synthetic model's Hellinger distance to ϕ is comparable to its relative L^2 norm by its bounded ratio.

We now estimate the entropy pairing for the normalized Hessian entries. Let $r_i = v_i/w_i$. The raw singleton Hessian has only diagonal entries because the other inputs are Gaussian. For its model,

$$w_i^2 \partial_i^2 L_i = \tfrac{1}{4}\mathcal{N}(\mathcal{N}-2)\mathcal{T}_{\rho w_i} g^\circ.$$

Using spare smoothing up to level $2b_i$ and (111) gives $\|w_i^2\partial_i^2 L_i\|_2 \le CD_i^b$. We compare the actual and model pairings

$$\int (w_i^2 \partial_i^2 \ell_i) \log(p/\phi) \quad \text{and} \quad \langle w_i^2 \partial_i^2 L_i, L\rangle_\phi.$$

The part from the model tested against $\log(p/\phi)-L$ is $O(D_i^{b+e})$, by $\mathcal{R} \geq 2^{-m}\phi$, (109) and $I^a \leq CD_i^e$. The residual $w_i^2\partial_i^2 t_i$ tested against this same log difference costs $O(D_i^{2e})$ by (105).

It remains to compare that singleton residual against each linear model L_j. Put $b_{ij} = \rho\sqrt{w_i w_j} \leq b_i$. OU self-adjointness for signed measures and its semigroup law give the exact identity

$$\int (w_i^2\partial_i^2 t_i)L_j = \tfrac{1}{4}\int [\mathcal{N}(\mathcal{N}-2)\mathcal{T}_{b_{ij}}g^\circ](x)\Delta_{b_{ij}}(x)\,dx. \tag{112}$$

For verification transfer the gain derivative to the Gaussian-smoothed test function L_j. At evaluation the normalized second derivative on degree n multiplies by $n(n-2)/4$. The product of the two channels of levels $\rho w_i, \rho w_j$ equals two channels of level b_{ij}, and one is moved back to the signed density $\mu-\phi(1+g^\circ)$. These manipulations may also be done by differentiating the Gaussian kernel. All test derivatives have polynomial growth and the signed measure has finite moments, so the pairings are well defined even if the raw relative density is unbounded.

Let $f = \mathcal{N}(\mathcal{N}-2)\mathcal{T}_{b_{ij}}g^\circ$. Under ϕ, $\|f\|_2 \leq CD_i^b$, by the reserve up to $2b_{ij}$ and (111). Also $|f(x)| \leq C(1+|x|^4)\|g^\circ\|_\infty$ by differentiating the OU kernel in x and using $b_{ij} \leq \rho < 1/8$. In the measure $\mathcal{R}_{b_{ij}}$, the part where $\mathcal{R}_{b_{ij}} \leq C_0\phi$ therefore has the same L^2 bound as under Gaussian measure. On the complement, the weighted mass of $\mathcal{R}_{b_{ij}}$ is charged to its singleton approximation error, since the true density dominates the model there. This adds $O(I^a)$ to the norm. Thus $\|f\|_{L^2(\mathcal{R}_{b_{ij}})} \leq CD_i^b$, and (112) costs $O(D_i^{b+e})$.

For the mixed part, its pairing with the logarithm is $O(D_i^{b+2\beta})$ for the normalized ij entry, by (107), (108) and $I \leq CD_i^{2\beta_0}$. Use whichever index has larger $|r_i|$. Then $|r_i r_j| \leq r_i^2$, so these terms are controlled by the same weighted sum of errors as the singleton terms. All their exponents $b+e, 2e, b+2\beta$ exceed two. Since $D_i \leq Cq$ uniformly, the total error is $o(1)\sum_i r_i^2 D_i^2$.

The leading Gaussian pairing is positive. On degree $n \geq 3$, put $k = n/2$. The contribution of L contains the positive multiplier $(\rho W)^k$, and the multiplier of its normalized i-th diagonal derivative is $k(k-1)(\rho w_i)^k$. The model has no degrees below three by exact mass, mean and covariance correction. Thus with $L'' = \partial_v^2 L$,

$$\langle L'', L\rangle_\phi \geq c\sum_i r_i^2\|\mathcal{T}_{b_i}g^\circ\|_2^2 \geq c'\sum_i r_i^2 D_i^2.$$

(The factor in the squared norm is b_i^n.) The exact entropy Hessian is the negative of $\int(\partial_v^2 p)\log(p/\phi)$, minus the nonnegative integral $\int(\partial_v p)^2/p$. The reference Gaussian logarithm drops out by isotropy. The differentiation under the entropy integral at each positive weight is justified by the same conditional polynomial kernel estimates and quadratic log bounds. Decreasing the common signal interval absorbs the error in the first pairing and proves Theorem 20.1. Positivity of D_i for non-Gaussian X follows from Gaussian deconvolution.

20.6 A logarithmic loss at the critical power

Theorem 20.2. *Fix $m \geq 3$ and $d \geq 1$. There are finite $B > 0$ and $q_0, c > 0$, depending only on m, d, such that the bound in Theorem 20.1 holds when*

$$0 < q \leq q_0, \qquad w_- \geq q^2[\log(1/q)]^B.$$

Proof. We give the choices and the uniform error budget for the preceding proof when its powers vary with q. Set

$$L_q = \log(1/q), \qquad \tau = \frac{\log L_q}{L_q}, \qquad M = L_q, \quad \rho = M^{-1}.$$

Reduce the common q-range so $m\rho < 1/8$. The total pre-signal Mq lies in the Fisher convexity range. All kernel bounds used above are uniform in these choices: normalized derivatives of the gains are bounded and the spare variances stay bounded away from zero.

There is a useful logarithmic version of the Hellinger estimate at all small signals. For $D = \mathsf{H}(p_s, \phi)$,

$$D_{\mathrm{KL}}(p_s \| \phi) \le C D^2 (1 + \log(1/D))^2.$$

Indeed its centered integrand is bounded by the Hellinger squared integrand times $C(1+|x|^2)$. Split at radius $C_0 \log(1/D)$; beyond it the uniform exponential tails of p_s, ϕ pay the same bound for C_0 sufficiently large. Similarly, for any fixed polynomial weight $(1+|x|)^K$, the signed-slot inequality has the logarithmic form

$$\int (1+|x|)^K \frac{(\mu_1 - \phi)^2}{(\mu_1 + \phi)/2} \le C_K I(\mu_1)(1 + \log(1/I(\mu_1)))^K. \tag{113}$$

Use the unweighted bound and $(\mu_1 - \phi)^2/((\mu_1+\phi)/2) \le 2(\mu_1 + \phi)$, split at radius $C_K \log(1/I(\mu_1))$, and use the uniform exponential tails on the exterior. These bounds also apply when μ_1 is one of the individual pre-smoothed inputs.

By assumption $s_i = q\sqrt{W w_i} \ge m^{-1/2} q^2 L_q^{B/2}$, so $\log(1/s_i) \le 2L_q - (B/2)\log L_q + O_m(1)$. Use the same modified-time interpolation from $3s_i/4$ to fixed s_h. For the wide pre-signal Mq, the coefficient θ on the low signal is at least

$$\frac{L_q - \log L_q - C}{2L_q - (B/2)\log L_q + C}.$$

In particular $2\theta \ge 1 + (B/4 - 1 - o(1))\tau$. For the narrow pre-signal Mqw_i, compare it by monotonicity to Ms_i. For the latter and for $2s_i$, the low coefficient is at least $1 - C\tau$, because the difference between the two lower times is $O(\log L_q)$ and the denominator is at least $L_q - O(1)$ in signal logarithms. The normalization I/t again costs a bounded factor since $\Theta(u) - u$ is bounded uniformly.

Here $D_i \le Cq$, so any fixed power of $1 + \log(1/D_i)$ is at most $C D_i^{-C'\tau}$. To see this, on $x = \log(1/D_i) \ge L_q - O(1)$, $\log(1+x)/x \le C \log L_q / L_q$. Combine the interpolation with the logarithmic KL estimate and entropy dissipation. For any prescribed $A > 0$, taking B large enough gives, uniformly,

$$I \le C D_i^{1+A\tau}, \qquad I(p_{Mqw_i}) + I(p_{2s_i}) \le C D_i^{2 - C_0\tau}. \tag{114}$$

The constant C_0 is independent of A, B. All statements are at sufficiently small q, which may depend on these chosen fixed constants.

Take $\gamma = 1 - K_0\tau$ in the truncation, with $K_0 > d/2 + 2$ fixed. Since $I \le C(Mq)^2$, we have $\log(1/I) \ge L_q$ for small q. Then the bound controlling the relative ratio on the ball, $I^{1-\gamma}(1+R)^d$, tends to zero uniformly. In fact writing $T = \log(1/I)$, this is at most $C e^{-K_0\tau T}(1+T)^{d/2}$, whose supremum for $T \ge L_q$ tends to zero. The covariance correction and (103) are unchanged, as γ stays above $1/2$. The intermediate-tail exponent in (105) is at least $\gamma/(1 + 2/M) \ge 1 - C\tau$. The far-tail exponent stays strictly above one. Powers of $1 + R$ are bounded by $I^{-C\tau}$ by the same elementary log bound now in T. The compact moment correction has exponent $\gamma = 1 - K_0\tau$. It follows that all singleton errors and the zero-order subset errors are $O(I^{1 - C_1\tau})$ in their weighted reciprocal norms, including derivatives of the singleton. For the zero-order products of at least two signed slots use (113) with the same I. The constant C_1 depends on fixed derivative and polynomial orders and on m, d, K_0, independently of A, B.

Consequently those errors are bounded by $C D_i^{1+(A-C)\tau}$ for every i, by (114). For the mixed normalized derivative use (113) in the individual pre-smoothed slots as in (107). One signed slot

uses the narrow exponent $2-C_0\tau$; at least one other uses the wide exponent $1+A\tau$. We may replace the logarithmic factors of a Fisher information by factors of D_i, since $u(1+\log(1/u))^K$ is increasing for sufficiently small u. This bounds the norm by $CD_i^{3/2+(A/2-C)\tau}$. Here and in the singleton bound the constants in the loss $C\tau$ can be chosen independently of A, B; when multiplying powers it is enough to take q so small that $A\tau \le 1/2$.

The logarithm bounds are still $C\sqrt{I}$ and $CI^{1-C_1\tau}$, respectively, in (108)–(109). The synthetic model upper estimate (111) and all the reserve-smoothed tests in (112) now cost $CD_i^{1-C\tau}$, by the narrow bound in (114) and the zero-order synthetic approximation. Thus the singleton comparisons, including the residual paired with the logarithm defect, have exponents at least $2+(A-C)\tau$, after allowing a larger constant C. The mixed derivative paired with the actual logarithm also has exponent at least $2+(A-C)\tau$, because the logarithm brings the additional $D_i^{1/2+A\tau/2}$. Choose A larger than all these fixed loss constants, and then B. Since $D_i^{c\tau} \le CL_q^{-c/2}$ for every fixed $c>0$, the errors are $o(1)\sum_i r_i^2 D_i^2$. The same margin ensures that the synthetic Hellinger approximation error is $o(D_i)$, so (110) and the leading Gaussian energy bound remain valid. This proves the theorem. □

21 Oscillator obstructions at unequal weights

21.1 Result

For independent real random variables X_1, X_2 with the same density $\gamma_1(1+g)$, and independent standard Gaussian G, write

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

We use even, bounded relative perturbations g for which the density has Gaussian mass and variance. Denote by $W(q)$ the supremum of $w \in [0,1/2]$ for which some positive relative supremum-norm radius guarantees strictly negative curvature for all nonzero such inputs whenever $\min(t,1-t) \le w$. One-sided derivatives are used at endpoint weights. The upper bound below also applies to the version with a common radius across dimensions, since the witnesses are scalar.

Theorem 21.1. *Fix $0<q<1$ and $0<t<1/2$. Set*

$$r = q(1-t), \qquad s = qt, \qquad d_r = \sqrt{1-r^2}, \qquad \lambda = \frac{r}{1+d_r}, \qquad \tau(b) = \frac{d_r - 1 + b}{1 + d_r - b}.$$

If $\tau(q)^{1/t} > \lambda$, there is a sequence of real, smooth, even relative perturbations converging to zero in every fixed relative C^k norm such that the densities are strongly log-concave and $F''_{q,g}(t) > 0$. Each density can match any prescribed finite set of Gaussian moments. At each fixed unequal weight this criterion holds for some signals strictly below the previous Fourier threshold $q = (1-t)/(1-t+t^2)$. Moreover,

$$W(q) \le \sqrt{1-q} - \frac{7}{12}(1-q) + O((1-q)^{3/2}) \qquad (q \uparrow 1).$$

The prior Fourier obstruction, the matching leading-order bound, and the dimension-uniform lower width $\sqrt{1-q}/(1+\sqrt{1-q})$ appear in Sections 13 and 16. Luo's Gaussian oscillator comparison and his midpoint discussion appear in [7, Problem 10.3 and Appendix F]. Evidence Press gives a positive-curvature midpoint profile for every $q > 4/7$ [17, Theorem B]. We treat prescribed *unequal* weights. The aligned oscillator norm comparison in Archive edition 11, §19.4, is an additional precursor. The sign in our result follows from a differentiated angular profile.

21.2 Replica Gram and separated product scales

Put $a = r$, $h_a = d_r$, $B = 1/h_a$ and $r_0 = \lambda^2$. Write T_b for the Gaussian Ornstein–Uhlenbeck operator with multiplier $b^{n/2}$ in Hermite degree n. Let $M(g_1, g_2)$ be the conditional product of $g_1(Z_1)$ and $g_2(Z_2)$ given $\sqrt{r}Z_1 + \sqrt{s}Z_2 + \sqrt{1-q}\,G$, for independent standard Gaussians. All output norms below use standard Gaussian measure.

For the oscillator basis of frequency B, use the conditional Gram of Section 11.3. We record the matrix properties needed here. Write $c = (\sqrt{r/q}, \sqrt{s/q})$, $\nu(q) = (1+q-h_a)/(1+q+h_a)$. For levels b_i with $\sum_i qc_i^2/b_i = 1$, put $D_\tau = \mathrm{diag}(\tau(b_1), \tau(b_2))$. The two matrices in the normalized Gram are

$$P = D_\tau^{-1/2}\Big[r_0 I + (\nu(q) - r_0)cc^\top\Big]\, D_\tau^{-1/2}, \quad Q = D_\tau^{-1/2}\Big[-r_0 I + (\tau(q) + r_0)cc^\top\Big]\, D_\tau^{-1/2}.$$

For $b_1 = b_2 = q$, P has eigenvalues $\nu(q)/\tau(q)$ and $r_0/\tau(q)$ in $(0, 1)$, and Q has eigenvalues 1 and $-r_0/\tau(q)$. The positive spectral margins persist on a small neighborhood of this allocation. The identity $\tau(b) + r_0 = 2h_a b/[(1 + h_a)(1 + h_a - b)]$ and the allocation relation give $I - Q \succeq 0$ with exactly one null direction along this neighborhood. In the physical replica form, the sum and difference have covariances $(I + P)/(2B)$ and $(I - Q)/(2B)$. The common collective variable is the null coordinate of the difference. The rank-one cross covariance makes the transverse replicas independent conditioned on it. All positive covariance eigenvalues are below $1/B$. Hence dividing out the oscillator ground states leaves a strictly decaying Gaussian on the replica support. Cauchy–Schwarz in the transverse pairing shows that the normalized Gram is bounded on Lebesgue L^2 of the two input coordinates. For input oscillator degrees m, n the attenuation in output norm is $\tau(b_1)^{m/2}\tau(b_2)^{n/2}$.

All constants may depend on the fixed q, t. Choose the sign of every even normalized oscillator H_j to give $H_j(0) > 0$. The exact even Hermite value at zero and Stirling's formula give $H_j(0) \asymp (j + 1)^{-1/4}$. Let $x = r/s > 1$. Given even $n \to \infty$, choose even m with $m = xn + O(1)$, so the physical frequencies $k_j = \sqrt{B(2j + 1)}$ satisfy $k_m\sqrt{s/q} - k_n\sqrt{r/q} = O(n^{-1/2})$.

At equal input smoothing levels $b_1 = b_2 = q$, the collective null vector of $I - Q$ is precisely $(\sqrt{r/q}, \sqrt{s/q})$. The aligned product with H_m on the heavy load and H_n on the light load has norm on exponential scale $\tau(q)^{(m+n)/2}$. Write

$$C = -\frac{x+1}{2}\log\tau(q), \qquad A = -\frac{1}{2}\log\lambda.$$

The hypothesis of Theorem 21.1 is $C < A$, since $x + 1 = 1/t$.

For the same-degree products, make a fixed small change of allocation u from $u = r/q$ to reduce $\tau(b_1)\tau(b_2)$. This is possible since at equal levels $(\log\tau(b_1)\tau(b_2))' = G(q)(1/s - 1/r) > 0$, where $G(b) = b^2\tau'(b)/\tau(b)$. Keep both levels strictly between a and 1. Set $E = -\frac{1}{2}\log(\tau(b_1)\tau(b_2))$ for that allocation, so $E > -\log\tau(q) = 2C/(x + 1)$. The product of two copies of H_n has norm and first two weight derivatives bounded by a polynomial times e^{-En}, and that of two copies of H_m has bound a polynomial times e^{-xEn}. For the reversed mixed product at the equal levels the exponent's allocation derivative is nonzero: the derivative of $n\log\tau(b_1) + m\log\tau(b_2)$, divided by n, tends to $G(q)(x/s - 1/r) > 0$. A small fixed decrease of the allocation gives a strict exponential gain in the upper bound. All these upper bounds include the first two actual weight derivatives at the target weight. These derivatives are finite sums of input Gaussian number operators of degree at most two in total. Each monomial sends an oscillator shell to uniformly shifted shells at polynomial cost. The bounded normalized replica Gram gives the desired upper bounds after adding back the corresponding attenuation factors.

The linear terms at both loads and their first two derivatives are bounded by a polynomial times e^{-An} or e^{-xAn}. At the larger reference load $a = r$ this follows from the diagonal oscillator Gram

and an upward signal change $a\exp(c_0/(n+1))$, which pays any fixed Gaussian number power at polynomial cost. The smaller load follows by contraction and similarly paid derivatives.

Take a low-mode coefficient $e^{-\eta n}$ and a high-mode coefficient one. The inequalities

$$\max(0, C-E) < \eta < \min(xE-C, xA-C) \tag{2}$$

are compatible. Indeed $xE > C$, $xA > C$, $2C < (x+1)E$, and $2C < xA+E$. The last inequality follows from $A > C$, $E > 2C/(x+1)$, and $x + 2/(x+1) > 2$ for $x > 1$. Consequently the aligned product dominates all other quadratic products on a strict exponential scale. In addition, both linear terms have strictly faster decay than the aligned product $e^{-(C+\eta)n}$. Multiplying the input by an extra $e^{-\epsilon n}$ for sufficiently small fixed $\epsilon > 0$ preserves this comparison after the linear and quadratic terms acquire their different powers of that factor. It gives convergence to zero in every fixed relative C^k norm and strict log-concavity. Derivative sup norms of the oscillator are polynomial by the ladder identity and one-dimensional Sobolev inequality.

To match the prescribed moments, replace each H_j by H_j minus a linear combination of $x^{2i}e^{-x^2}$, up to the prescribed order. The weighted polynomial Gram matrix is invertible. The Gaussian moment generating coefficient of H_j is a fixed polynomial times $\exp(\lambda^2 z^2/2)$ in the oscillator generating variable z. Stirling's formula bounds each correction coefficient by a polynomial times λ^j. The corrections have Gaussian L^2, sup and number-domain sizes bounded by a polynomial times λ^m, λ^n, respectively. With the above amplitudes an output error containing one correction, including its first two derivatives, is bounded by a polynomial times $e^{-(2A+\eta)n}$ or e^{-2xAn}; this follows from Gaussian tensor contraction and positive output smoothing on a fixed weight neighborhood. These rates both beat $e^{-(C+\eta)n}$, by (2) and $C < A$. Linear correction errors are smaller as well.

21.3 Signed alignment at the prescribed weight

It remains to derive the curvature sign, which does not follow from the norm rate alone. Keep $b_1 = b_2 = q$ as the physical weight varies. Write its smaller weight as $\sin^2\alpha$ locally and let α_0 denote the target angle. The normalized Gram diagonalizes in the moving collective and transverse directions $e = (\cos\alpha, \sin\alpha)$, $e_\perp = (-\sin\alpha, \cos\alpha)$. Its eigenvalues and determinant are independent of α. In the physical replica representation above, the common collective coordinate w is independent of the transverse vector of either replica. The two transverse coordinates are independent given w, and their conditional distribution is therefore independent of w and α. Undoing the oscillator ground factors gives the exact representation

$$\|M_{q,\alpha}(H_m, H_n)\|_2^2 = K_0\tau(q)^{m+n}\int j(w)|J_\alpha(w)|^2\,dw, \tag{3}$$

where

$$J_\alpha(w) = \int H_m(w\cos\alpha - z\sin\alpha)H_n(w\sin\alpha + z\cos\alpha)\phi(z)\,dz.$$

Here $K_0 > 0$, $j(w) > 0$ is a fixed decaying Gaussian, and $\phi(z) > 0$ is a centered fixed Gaussian with variance $\sigma^2 > 0$ after normalization. The strict covariance gaps that yield the densities hold at equal levels $q \in (a, 1)$.

Let $\alpha = \alpha_0 + y/\sqrt{n}$. Divide the inner integral in (3) by $H_m(0)H_n(0)$. For y in a fixed compact interval, replace each oscillator ratio by its cosine approximation, with an error tending to zero in the weighted $C_y^2 L_w^2(j\,dw)$ norm. To justify derivatives, on $|u| \le R$, let $f_j(u) = H_j(u)/H_j(0)$. The oscillator solves $H_j'' + k_j^2 H_j = B^2x^2H_j$ with $H_j'(0) = 0$. Its sine-kernel Duhamel equation gives the difference $f_j(u) - \cos(k_j u)$ and its first derivative divided by k_j as $O(R^3/k_j)$ uniformly for $R = C_1\sqrt{\log(n+2)}$. The oscillator equation gives the same for the second derivative divided by

k_j^2, up to $O(R^2/k_j^2)$. The derivatives in y of either oscillator argument are $O((|w|+|z|)/\sqrt{n})$ and $O((|w|+|z|)/n)$. The resulting polynomial factors in R still leave an $o(1)$ error. Use $|w|+|z| \le R$ and then Gaussian tails to control the complement, with the polynomial global bounds for the ratios and their first two derivatives.

The sum phase has collective and transverse frequencies

$$K_s(\alpha) = k_m \cos\alpha + k_n \sin\alpha, \qquad K_z(\alpha) = -k_m \sin\alpha + k_n \cos\alpha.$$

Set $K = K_s(\alpha_0)$, so $K/\sqrt{n} \to \kappa = \sqrt{2B(x+1)}$. On the y scale, $K_s - K \to 0$ and $K_z \to -\kappa y$ in C^2 on compact intervals. The difference phase has transverse frequency of order $\sqrt{n}$, so its Gaussian transform vanishes with the same derivative control. Direct Gaussian integration of the sum phase gives, in the above weighted C^2 sense,

$$\frac{1}{2}\left(\int \phi\right) e^{-\sigma^2\kappa^2 y^2/2} \cos(Kw) + o(1).$$

Since $K \to \infty$, $\int j(w)\cos^2(Kw)\,dw \to \frac{1}{2}\int j(w)\,dw$. It follows from (3) that the squared product norm divided by $\tau(q)^{m+n}H_m(0)^2H_n(0)^2$ tends in C_y^2 to a positive constant times $e^{-\sigma^2\kappa^2y^2}$. Its second derivative with respect to the physical weight at the target is strictly negative of order n times the norm squared. Indeed the y-derivative term from differentiating the angle-to-weight change contributes only $O(\sqrt{n})$ in this normalization.

21.4 Entropy curvature

Let $R = L + M$ be the output Gaussian relative perturbation for the scaled, corrected input. The pointwise bound $\|R\|_\infty$ is exponentially small by conditional contraction. The quadratic entropy term is $-\frac{1}{2}\|R\|_2^2$; exactly twice differentiating the entropy, its difference from this quadratic term in curvature is bounded by

$$C\|R\|_\infty(\|R\|_2\|R''\|_2 + \|R'\|_2^2).$$

This follows from $\log(1+R) - R = O(R^2)$ and $(1+R)^{-1} - 1 = O(R)$. All norms and derivatives are bounded by polynomial factors times the aligned exponential scale, while nonaligned output terms are exponentially smaller with two derivatives by (2). Formula (3) gives the main squared-norm curvature with a nonzero negative coefficient, so the entropy curvature is strictly positive.

21.5 Threshold and high-signal width

Set $c = 1 - q$. Since $\tau(q) = (h_a - c)/(h_a + c)$ and $\lambda = \sqrt{(1-h_a)/(1+h_a)}$, the criterion of Theorem 21.1 is

$$2\,\mathrm{artanh}(c/h_a) < t\,\mathrm{artanh}(h_a). \tag{4}$$

Using $\mathrm{artanh}(h_a) = 2\,\mathrm{artanh}((1-r)/h_a)$, this is a strict improvement at every old Fourier threshold point $c = t(1-r)$, by strict convexity of artanh on $(0,1)$. It therefore reaches strictly lower signals at each fixed unequal weight.

For any fixed q, the right side of (4) strictly increases with t on $(0, 1/2]$, because h_a does, while the left side strictly decreases. At $t = 0$ the inequality is false, and at the balanced point its strict form is $\tau(q)^2 > \lambda$. At $a = q/2$, clearing denominators in $\tau(q)^2 > \lambda$ gives $7q > 4$, as in Section 11.4. Thus for every $q > 4/7$ there is a unique $t_{\mathrm{osc}}(q) \in (0, 1/2)$ satisfying equality in (4). Every unequal nearer-endpoint weight above that threshold admits the obstruction. This supplies a closed implicit upper bound for the full uniform negative-curvature half-width at each such q.

For the curve of equality near $c = t = 0$, use $h_a^2 = 1 - (1-c)^2(1-t)^2$ and put $c = t^2 v$. Multiplying (4) by h_a and dividing the equality by $2t^2$ gives an analytic function of v, t near $(1, 0)$, by the power series for $\operatorname{artanh} u/u$. At $t = 0$ the equality is $v = 1$ with nonzero derivative. Expanding $h_a^2 = 2t + t^2 + O(t^3)$ on that branch gives

$$2c + O(t^5) = t(h_a^2 + h_a^4/3 + O(h_a^6)) = 2t^2 + \tfrac{7}{3}t^3 + O(t^4).$$

Thus $c = t^2 + \frac{7}{6}t^3 + O(t^4)$, and with $\beta = \sqrt{c}$ the qualifying upper width is $t = \beta - \frac{7}{12}\beta^2 + O(\beta^3)$. The implicit branch has the correct inequality direction by its nonzero derivative. This improves the earlier Fourier upper coefficient $-1/2$ while leaving a gap to the binomial lower coefficient -1.

21.6 Two oscillator widths

The angular construction extends when the two modes have different oscillator precisions. The equal-precision criterion in Theorem 21.1 gives an explicit sufficient condition at every unequal weight. In the high-signal regime the two-precision variant sharpens the next term of the resulting upper bound.

Theorem 21.2. *Set $\beta = \sqrt{1-q}$. For every fixed $0 < \kappa_1 < 1/2$,*

$$W(q) \le \beta - \frac{7}{12}\beta^2 + \left(\frac{769}{1440} - \frac{\kappa_1^2}{3}\right)\beta^3 + O_{\kappa_1}(\beta^4) \qquad (q \uparrow 1).$$

In particular,

$$\limsup_{q\uparrow 1} \frac{W(q) - \beta + \frac{7}{12}\beta^2}{\beta^3} \le \frac{649}{1440}.$$

21.6.1 Gaussian replica calculation for two widths

Use ordered loads $r = q(1-t), s = qt$, write $c = 1 - q$ and $\ell = (\sqrt{r}, \sqrt{s})^\top$. Use scalar oscillator precisions $B_1, B_2 > 1$, $B = \operatorname{diag}(B_1, B_2)$. For input oscillator degree n_i the physical oscillator equals $B_i^{1/4} H_{n_i}(\sqrt{B_i}\, x)$ with the unit-frequency normalization. Integrating the two oscillator generating functions in two conditionally independent Gaussian replicas, with covariance $I \pm \ell\ell^\top$ in the sum/difference coordinates, gives matrices before attenuation

$$P_0 = 2\sqrt{B}\big((I + \ell\ell^\top)^{-1} + B\big)^{-1}\sqrt{B} - I, \quad Q_0 = I - 2\sqrt{B}\big((I - \ell\ell^\top)^{-1} + B\big)^{-1}\sqrt{B}.$$

Their Bargmann exponent has the same form as the normalized Gram in Section 2. Define

$$\tau_i(b) = \frac{1 - B_i(1-b)}{1 + B_i(1-b)}, \quad p_i = \frac{B_i - 1}{B_i + 1}, \quad v_i = \frac{\sqrt{B_i}\,\ell_i}{1 + B_i}.$$

The Sherman–Morrison formula yields

$$P_0 = \operatorname{diag}(p_i) + \frac{2vv^\top}{1 + q - \sum_i \ell_i^2/(1+B_i)}, \qquad Q_0 = -\operatorname{diag}(p_i) + \frac{2vv^\top}{1 - q + \sum_i \ell_i^2/(1+B_i)}.$$

For allocations $\sum_i \ell_i^2/b_i = 1$, conjugate each by $D_\tau^{-1/2}$ with diagonal $\tau_i(b_i)$. The identity $\tau_i(b_i) + p_i = 2B_i b_i/[(1+B_i)(1+B_i(1-b_i))]$ gives $I - Q$ nonnegative with a single null direction, as its rank-one condition is exactly the allocation equality. The positive unit null vector e has

$$e_i^2 \propto \frac{\ell_i^2[1 - B_i^2(1-b_i)^2]}{B_i b_i^2}. \tag{5}$$

At equal levels $b_i = q$, suppose $B_i^2(1-q^2) < 1$ for both widths. Then $\tau_i(q) > p_i > 0$ and $Q > -I$. Also $P > 0$. For $P < I$ the rank-one comparison is a weighted average, with weights ℓ_i^2/q, of the comparisons for both widths separately treated as homogeneous. For a homogeneous width B_i it is equivalent to $\nu_i(q) < \tau_i(q)$, where $\nu_i(q) = [B_i(1+q)-1]/[B_i(1+q)+1]$. This follows from $B_i^2(1-q^2) < 1$. Thus all covariance gaps are strict at equal levels and persist for small level perturbations that preserve the allocation. The physical normalized Gram is represented in dimensionless oscillator variables by replica sum/difference covariances $(I+P)/2, (I-Q)/2$. The two replica projections onto e coincide. Conditional on their common value, the transverse variables are independent. All positive covariance eigenvalues are below one. After stripping the oscillator ground states the joint Gaussian decays strictly. In particular the normalized Gram gives the same $L^2(dx)$ operator upper bound as before, with output attenuation $\tau_1(b_1)^{n_1/2}\tau_2(b_2)^{n_2/2}$.

21.6.2 Stationarity and comparison rates

Choose $B_2 = (1-r^2)^{-1/2}$ and $B_1 = \kappa_1/t$. We use $c = t^2 + O(t^3)$, so the spectral hypotheses just proved hold for all sufficiently small $t > 0$. The relation $B_1^2(1-q^2) \to 2\kappa_1^2 < 1$ and $B_2^2(1-q^2) \to 0$ suffice. At equal levels q, write $z_i = B_i c$, $U = \operatorname{artanh} z_1$, $V = \operatorname{artanh} z_2$. Take high degree $m = xn + O(1)$, with even rounding and

$$x = \frac{r}{s}\frac{B_2}{B_1}\frac{1-z_1^2}{1-z_2^2}, \qquad C = xU + V.$$

Equation (5) gives $m/n \to e_1^2/e_2^2 = x$. The main aligned product has norm on exponential scale e^{-Cn}.

For a width B_i, the scalar linear norm at the heavier load r, including a fixed number of derivatives, has rate at least $A_i = -\frac{1}{2}\log\max\{\nu_i(r), \tau_i(r)\}$. This follows from the generating diagonal $((1-\nu_i(r)z)(1-\tau_i(r)z))^{-1/2}$ and a small upward change in r paying number powers. The bases here have both parameters positive for small t. Consequently

$$A_2 = \tfrac{1}{2}\operatorname{artanh}\sqrt{1-r^2}, \qquad A_1 = \operatorname{artanh}\frac{1}{B_1(1+r)}.$$

For the self-products at the two widths, tilt the allocation slightly to give decay rates $E_1 > 2U$ per high index and $E_2 > 2V$ per low index. For either width the derivative of the log product of the two levels at $u = r/q$ equals $G_i(q)(1/s - 1/r) > 0$, where $G_i(b) = b^2\tau_i'(b)/\tau_i(b)$. Indeed $d\log\tau_i(b)/db > 0$ and the two loads are unequal. For the reverse mixed pair, stationarity under allocations would require $n/m = (r/s)(B_1/B_2)(1-z_2^2)/(1-z_1^2)$, which contradicts the displayed value of x since $r > s$. A tilt gives strict exponential improvement over the main rate. These allocation gains persist when paying two actual weight derivatives, since number operators shift oscillator levels by bounded amounts at polynomial cost.

If $C < A_2$, take low amplitude $\exp(-\eta n)$ and high amplitude one. The needed bounds are

$$\max(0, C - E_2) < \eta < \min(xE_1 - C, xA_1 - C).$$

They are compatible for small t on the stated high-signal scale. Indeed $xU \asymp \sqrt{t}$, $V = O(t^{3/2})$, so $xE_1 - C > 0$. Also $2C = 2xU + 2V < xE_1 + E_2$. Finally $xA_1/C \to 1/(2\kappa_1^2) > 2$, so $xA_1 - C$ exceeds $\max(0, C - E_2)$. The same exponential suppression of moment corrections holds because their rate at width B_i, from the Gaussian oscillator generating integral, is $\operatorname{artanh}(1/B_i) > A_i$; for B_2 it is $2A_2$. An overall sufficiently small extra exponential gives the same strongly log-concave relative C^k witnesses.

21.6.3 Sign and expansion

In the scaled angle proof of Section 3, use the angle of the normalized collective vector e instead of the angle of the underlying loads. Freeze B_1, B_2 at the prescribed weight when differentiating. At equal levels q both attenuation factors then remain constant as the physical mixing weight varies. Equation (5) makes the e angle strictly monotone in that weight. The normalized Gaussian densities and transverse conditional mean now vary smoothly with the weight. On a $n^{-1/2}$ variation of the weight they converge in rescaled C^2 to the fixed density parameters. Conditional Gaussian integration of the aligned cosine sum gives a damping factor $e^{-\sigma^2 K_z^2/2}$ and a phase shift $\mu K_z w$ from the conditional transverse mean μw. The integrated cosine square splits into its constant term and a Gaussian Fourier transform evaluated at twice the shifted collective frequency. That frequency tends to infinity, while its first two scaled derivatives stay bounded on compact scaled intervals. The Gaussian transform and its derivatives vanish uniformly there. The Gaussian tails and polynomial oscillator bounds pay the derivatives as before. The limiting squared norm is thus a positive Gaussian in the scaled weight difference, with strictly negative second derivative at the target. The entropy Hessian remainder and amplitude arguments are unchanged.

For the sharper width, write $h = \sqrt{1 - r^2}$. The equality $C = A_2$ gives, using the power series for artanh,

$$\frac{C}{B_2} = \frac{c}{t}\left(1 - \frac{2}{3}\kappa_1^2 t^2 + O(t^3)\right), \qquad \frac{A_2}{B_2} = \frac{1}{2} h \operatorname{artanh} h.$$

To see the first expansion, $z_1^2 = \kappa_1^2 t^2 + O(t^3)$, $z_2^2 = O(t^3)$, and $(1-z^2)\operatorname{artanh}(z)/z = 1 - \frac{2}{3}z^2 + O(z^4)$. The analytic equality branch with $c/t^2 \to 1$ therefore has

$$c = t^2 + \frac{7}{6}t^3 + \left(\frac{19}{30} + \frac{2}{3}\kappa_1^2\right) t^4 + O_{\kappa_1}(t^5).$$

For the 19/30 term, substitute $c = t^2 + (7/6)t^3 + O(t^4)$ in $h^2 = 1 - (1-c)^2(1-t)^2 = 2t + t^2 - \frac{5}{3}t^3 + O(t^4)$, then expand $t(h^2 + h^4/3 + h^6/5)/2$. The implicit branch is analytic after writing $c = t^2 v$, since the normalized equation contains only analytic functions of z_i^2 and h^2, with $z_2^2 = t^3$ times an analytic factor. Its derivative in v is nonzero at $t = 0, v = 1$. The inverse function and strict side give the upper bound in Theorem 21.2. Taking $\kappa_1 \uparrow 1/2$ gives the limsup in Theorem 21.2. These two-width witnesses improve the upper bound at third order. Matching the second coefficient requires a further mechanism or a stronger lower bound.

22 The first unmatched Gaussian cumulant of a log-concave law

Let γ denote the standard Gaussian density on $\mathbb{R}$. Write $\mathcal{F}_n$ for the log-concave probability densities that match its moments in orders $0, \dots, n-1$, where $n \geq 3$. All such laws are centered and have variance one. Put

$$A_n = \sup_{f \in \mathcal{F}_n} \frac{|\kappa_n(f)|}{\sqrt{n!}}.$$

The normalization is the orthonormal Hermite coefficient at the first degree whose moment is free. For $0 < \delta < 1$ also put

$$A_n^{\mathrm{rel}}(\delta) = \sup\left\{\frac{|\kappa_n(f)|}{\sqrt{n!}} : f \in \mathcal{F}_n, \quad \left\|\frac{f}{\gamma} - 1\right\|_\infty \leq \delta\right\}.$$

Theorem 22.1. *There are absolute constants $c, C > 0$ such that for every $n \geq 3$,*

$$c(n+1)^{-3/4} 2^{-n/2} \leq A_n \leq C(n+1)^{1/4} 2^{-n/2}.$$

For every fixed $0 < \delta_0 < 1$, there are $c, C_{\delta_0} > 0$ such that for every $n \geq 3$ and $0 < \delta \leq \delta_0$,

$$c\,(n+1)^{1/4} 2^{-n/2} \min\{\delta, (n+1)^{-1}\} \leq A_n^{\mathrm{rel}}(\delta) \leq C_{\delta_0} (n+1)^{1/4} 2^{-n/2} \min\{\delta, (n+1)^{-1}\}.$$

The lower estimates hold for each sign of the cumulant separately, with smooth densities satisfying $(-\log f)'' \geq 1/2$.

The bounds apply to the finite moment matching in the moment-matched constructions of Sections 7 and 17 and in Luo's related entropy counterexamples [7]. They quantify how large the first permitted cumulant can be as the matching order increases. The earlier analyses use moment matching to remove prescribed lower terms and study the curvature at given smoothing and mixing weights. The bound here concerns the cumulant itself.

22.1 The upper bound over all log-concave laws

Let He_j be the probabilists' Hermite polynomial, with leading coefficient one. The Gaussian inner products satisfy $\int \mathrm{He}_i \mathrm{He}_j \, \gamma = j! \, \mathbf{1}_{i=j}$, and the recurrence is $\mathrm{He}_j(x) = x \, \mathrm{He}_{j-1}(x) - (j-1) \, \mathrm{He}_{j-2}(x)$ for $j \geq 2$. The roots of He_j are the eigenvalues of the j-by-j real symmetric tridiagonal matrix with zero diagonal and off-diagonal entries $\sqrt{1}, \ldots, \sqrt{j-1}$, by the determinant recurrence. In particular they are real and have absolute value at most $2\sqrt{j}$ by the maximum row sum bound. Beyond $2\sqrt{j}$ the polynomial is positive and increasing.

Lemma 22.2. *Suppose X is centered, has variance one and is log-concave. If $j \geq 2$, $\mathbb{E}\,\mathrm{He}_{j-1}(X)^2 = (j-1)!$ and $\mathbb{E}\,\mathrm{He}_{j-2}(X)^2 = (j-2)!$, then $\mathbb{E}\,\mathrm{He}_j(X)^2 \leq Cj!$ with a universal constant.*

Proof. Set $T = 24\sqrt{j}$ and let $S(x) = \mathbb{P}(X \geq x)$. On $\{|X| \leq T\}$, the recurrence and the L^2 triangle inequality bound the square root of the truncated expectation by

$$T\sqrt{(j-1)!} + (j-1)\sqrt{(j-2)!} \leq 25\sqrt{j!}.$$

By increasing positivity of He_{j-1} on $[T, \infty)$,

$$S(T) \leq \frac{(j-1)!}{\mathrm{He}_{j-1}(T)^2} \leq 22^{-2(j-1)}.$$

The second inequality follows from the root bound and $(j-1)! \leq j^{j-1}$. The survival function is log-concave [10], with $S(-2) \geq 3/4$ by Chebyshev. When $S(T) > 0$, the secant estimate gives $S(T+s) \leq S(T) e^{-\lambda s}$ for $s \geq 0$, where

$$\lambda = \frac{2(j-1)\log 22 - \log(4/3)}{T+2}.$$

On $[T, \infty)$ the logarithmic derivative of He_j^2 is at most $d = 2j/(T - 2\sqrt{j})$. The ratio λ/d is bounded below for $j \geq 2$ by

$$\frac{22}{24+\sqrt{2}} \left(\frac{\log 22}{2} - \frac{\log(4/3)}{4} \right) > \frac{5}{4}.$$

Integrating by parts against the survival function, or integrating the derivative from T to X on $\{X \geq T\}$, gives

$$\begin{aligned} \mathbb{E}[\mathrm{He}_j(X)^2 \mathbf{1}_{X \geq T}] &\leq \mathrm{He}_j(T)^2 S(T) \left(1 + \frac{d}{\lambda - d}\right) \\ &\leq 5T^2 (j-1)!. \end{aligned}$$

Here $\mathrm{He}_j(T) \leq T \, \mathrm{He}_{j-1}(T)$ by the recurrence. If $S(T) = 0$, the bound is immediate. Reflection handles the other tail. Adding the three estimates proves the lemma. □

Since the moments through order $n-1$ agree, $\kappa_n(f) = \int \mathrm{He}_n f$. If $n = 2k$, the polynomial $\mathrm{He}_k^2 - \mathrm{He}_n$ has degree at most $n-1$, and Gaussian orthogonality yields

$$\kappa_n(f) = \mathbb{E}_f \,\mathrm{He}_k^2 - k!.$$

Lemma 22.2 gives $|\kappa_n(f)| \le Ck!$. If $n = 2k+1$, the same argument gives $\kappa_n(f) = \mathbb{E}_f[\mathrm{He}_k \,\mathrm{He}_{k+1}]$. Since $\mathbb{E}_f \,\mathrm{He}_k^2 = k!$, Cauchy–Schwarz and the lemma give $|\kappa_n(f)| \le C\sqrt{k!(k+1)!}$. Stirling's bounds in both cases give

$$\frac{|\kappa_n(f)|}{\sqrt{n!}} \le C(n+1)^{1/4} 2^{-n/2}.$$

22.2 A bounded Hermite perturbation

For $j \ge 0$ let

$$u_j(y) = \frac{1}{\sqrt{j!}} e^{-y^2/4} \,\mathrm{He}_j(y), \qquad N_j = j + \tfrac{1}{2}.$$

The oscillator equation and ladder identity are

$$u_j'' + (N_j - y^2/4)u_j = 0, \qquad 2u_j' = \sqrt{j}\, u_{j-1} - \sqrt{j+1}\, u_{j+1},$$

where the term u_{j-1} is absent for $j = 0$. These identities follow by differentiating the Hermite recurrence and using $\mathrm{He}_j' = j \,\mathrm{He}_{j-1}$. For $0 \le y \le 2\sqrt{N_j}$, the derivative of the energy

$$u_j'(y)^2 + (N_j - y^2/4)u_j(y)^2$$

is $-y u_j(y)^2/2$. At zero the energy is $O(\sqrt{j+1})$. Indeed $u_{2k}(0)^2 = 4^{-k}\binom{2k}{k}$, $u_{2k}'(0) = u_{2k+1}(0) = 0$ by parity, and $u_{2k+1}'(0) = \sqrt{2k+1}\, u_{2k}(0)$. The energy therefore bounds $|u_j'|$ by $C(j+1)^{1/4}$ up to the turning point. Beyond that point u_j has constant sign. For example, a zero there would give a contradiction upon integrating the oscillator equation times u_j from that zero to infinity, since the potential $y^2/4 - N_j$ is nonnegative. The equation then shows that the derivative decreases in absolute value towards zero there. Parity treats the negative half-line. Thus $\|u_j'\|_\infty \le C(j+1)^{1/4}$. Differentiating the twice-applied ladder gives $\|u_j'''\|_\infty \le C(j+1)^{5/4}$.

Put

$$b_n(x) = \left(u_n - \sqrt{\frac{n+2}{n+1}}\, u_{n+2}\right)(\sqrt{2}x) = \frac{2}{\sqrt{n+1}} u_{n+1}'(\sqrt{2}x).$$

The estimates above give $\|b_n\|_\infty \le C(n+1)^{-1/4}$ and $\|b_n''\|_\infty \le C(n+1)^{3/4}$. Changing variables $y = \sqrt{2}x$ gives

$$\int x^i u_j(\sqrt{2}x)\gamma(x)\,dx = \frac{2^{-(i+1)/2}}{\sqrt{j!}} \int y^i \,\mathrm{He}_j(y)\gamma(y)\,dy.$$

This is zero if $i < j$, and equals $2^{-(j+1)/2}\sqrt{j!}$ if $i = j$. Thus b_n cancels exactly through degree $n-1$ and has the same degree-n moment as $u_n(\sqrt{2}x)$.

For $0 < \delta \le \delta_0$, choose

$$\epsilon = c_0 \min\{\delta(n+1)^{1/4}, (n+1)^{-3/4}\}, \qquad f = \gamma(1 \pm \epsilon b_n)$$

with a universal sufficiently small $c_0 > 0$. The moment equalities ensure normalization and the required matching. The relative perturbation has sup norm at most δ, and we can also make that norm at most $1/2$. Its second derivative is bounded in norm by an arbitrarily small constant by choosing c_0. Since $(\log(1+g))'' = g''/(1+g) - (g')^2/(1+g)^2$ for $g = \pm\epsilon b_n$, the log density has second derivative at most $-1/2$. The normalized cumulant is $\pm\epsilon 2^{-(n+1)/2}$. This proves both lower bounds in Theorem 22.1.

22.3 The sharp bound in a density-ratio neighborhood

We need two elementary integral bounds for the oscillator. For every $j \ge 0$,

$$\int |u_j(y)|\,dy \le C(j+1)^{1/4}, \qquad \sum_{k\in\mathbb{Z}} \sup_{k\le y\le k+1} |u_j(y)| \le C(j+1)^{1/4}. \tag{115}$$

For the first bound, Gaussian orthogonality gives $\|u_j\|_2^2 = \sqrt{2\pi}$. The recurrence gives $\|yu_j\|_2^2 \le C(j+1)$, so weighted Cauchy–Schwarz with weight $1 + y^2/(j+1)$ proves the estimate. For the second, a fixed number of cells adjoining the turning points cost at most $C(j+1)^{1/4}$, since $\|u_j\|_\infty^2 \le 2\|u_j\|_2\|u_j'\|_2$ and the ladder gives $\|u_j'\|_2 \le C\sqrt{j+1}$. For $0 \le y < 2\sqrt{N_j}$ the energy estimate yields

$$|u_j(y)| \le C(2\sqrt{N_j} - y)^{-1/2}.$$

Sum this over the interior cells at distance at least one from the turning point. The sum is $O((j+1)^{1/4})$. On the exterior half-line the absolute value is monotone decreasing, so the cell suprema are controlled by the L^1 estimate and at most one extra sup norm. Reflection completes the proof of (115). The same bounds hold, up to universal constants, for $u_j(\sqrt{2}x)$ integrated or summed over unit cells in x.

Now write $f = \gamma(1+g)$ with $\|g\|_\infty \le \delta$. The monic polynomial $2^{-n/2}\,\mathrm{He}_n(\sqrt{2}x)$ can replace $\mathrm{He}_n(x)$ in $\int g\,\mathrm{He}_n\,\gamma$ by moment matching. Its product with γ is a constant times $2^{-n/2}\sqrt{n!}u_n(\sqrt{2}x)$. Thus (115) and the bound on g give

$$|\kappa_n(f)| \le C\delta(n+1)^{1/4}2^{-n/2}\sqrt{n!}. \tag{116}$$

For the second bound, let $L = \log(1+g)$. Log-concavity of f gives $L'' \le 1$ distributionally. At every differentiability point, semiconcavity gives

$$L(y) \le L(x) + L'(x)(y-x) + \tfrac{1}{2}(y-x)^2.$$

Taking $y = x - L'(x)$ shows $|L'|^2 \le 2\operatorname{osc} L$ almost everywhere, with the oscillation bounded in terms of δ_0. The derivative of L has locally bounded variation, and hence

$$g'' = (1+g)L'' + (1+g)(L')^2\,dx$$

is a signed measure. Its positive part is bounded by $C_{\delta_0}\,dx$. Since $g' = (1+g)L'$ is uniformly bounded, the mass of $|g''|$ in every interval of length one is at most C_{δ_0}: the negative mass is bounded by the positive mass and the boundary increment of g'. The argument uses one-sided limits of the bounded-variation representative at the ends.

Twice integrating by parts with $(\mathrm{He}_j\,\gamma)' = -\,\mathrm{He}_{j+1}\,\gamma$ gives

$$\int \mathrm{He}_n\,\gamma g\,dx = \int \mathrm{He}_{n-2}\,\gamma\,dg''.$$

The identity is valid for the measure g'' by truncation, using its uniform local variation bound and Gaussian decay. The same identity with n replaced by $i+2 < n$ shows that $g''\gamma$ annihilates every polynomial of degree less than $n-2$. Replace He_{n-2} by the monic polynomial $2^{-(n-2)/2}\,\mathrm{He}_{n-2}(\sqrt{2}x)$ in the last pairing. Using the cell supremum estimate in (115) gives

$$\begin{aligned} |\kappa_n(f)| &\le C_{\delta_0}2^{-(n-2)/2}\sqrt{(n-2)!}\,(n-1)^{1/4} \\ &\le C_{\delta_0}(n+1)^{-3/4}2^{-n/2}\sqrt{n!}. \end{aligned}$$

Together with (116) this proves the required joint bound in the radius and the degree.

22.4 Directional cumulants and Gaussian smoothing

Corollary 22.3. *The upper and lower orders of Theorem 22.1 hold in every dimension for the supremum of the normalized absolute directional cumulant over unit vectors. The law is now required to match all standard Gaussian mixed moments through degree $n-1$, and the density ratio is taken relative to the standard Gaussian in that dimension. The constants are independent of the dimension.*

Proof. Every unit projection is log-concave, by integration over the orthogonal coordinates, and matches the required moments by the multinomial expansion. The density ratio of that projection is the conditional Gaussian average of the original density ratio, so it obeys the same relative bound. Apply the one-dimensional estimate. Products of the one-dimensional witnesses with independent Gaussian coordinates give the lower bounds, including the strong log-concavity assertion. □

For the entropy application, let X have density $f \in \mathcal{F}_n$ with $\|f/\gamma - 1\|_\infty \le \delta$ and let G be independent standard Gaussian. Write $X_q = \sqrt{q}X + \sqrt{1-q}G$ and denote the density of X_q by f_q. Its first free Hermite coefficient is $q^{n/2}\kappa_n(f)/\sqrt{n!}$ because the Ornstein–Uhlenbeck operator multiplies the jth Hermite level by $q^{j/2}$. For fixed n and $\delta \le \delta_0 < 1$,

$$\lim_{q\downarrow 0} q^{-n} \sup_f D(f_q\|\gamma) = \tfrac{1}{2}(A_n^{\mathrm{rel}}(\delta))^2,$$

where the supremum has the same constraints as $A_n^{\mathrm{rel}}(\delta)$. Indeed, the squared relative $L^2(\gamma)$ norm of $f_q/\gamma - 1$ is $q^n\kappa_n(f)^2/n! + O(\delta^2 q^{n+1})$. The error from the quadratic expansion of relative entropy is bounded by $C_{\delta_0}\|f_q/\gamma - 1\|^3_{L^3(\gamma)}$. For $q \le 1/2$, Gaussian hypercontractivity from L^2 to L^3 bounds the latter by $C_{\delta_0}\delta^3(2q)^{3n/2}$. Both errors are uniform over the indicated fiber. This gives the largest possible leading entropy deficit in the bounded-relative moment-matched class.

23 Fixed asymmetric inputs and the endpoint transition of mixing entropy

23.1 Eventual concavity for each fixed law

Use the Gaussian-noised mixing entropy

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

where the X_i are independent copies of X and G is independent standard normal. Let X be isotropic in $\mathbb{R}^d$. Suppose that its first moment mismatch from the standard Gaussian occurs at total degree $n \ge 3$, and define

$$\mathcal{P}(x) = \sum_{|\nu|=n} \frac{\mathbb{E}X^\nu - \mathbb{E}G^\nu}{\nu!}\,\mathrm{He}_\nu(x), \qquad k = n/2.$$

Here the Hermites are the product probabilists' Hermite polynomials. Put $S_k(w) = \sum_i w_i^k$, $D_k(w,v) = \sum_i v_i^2 w_i^{k-2}$, and $U_k(w,v) = \sum_i v_i w_i^{k-1}$. On the relative interior of a face these sums use its positive weights and directions tangent to it.

Theorem 23.1. *If $\mathbb{E}|X|^{2n+\epsilon} < \infty$ for some $\epsilon > 0$, then, uniformly on the relative interiors of all faces and for tangents to the face as $q \downarrow 0$,*

$$\partial_v^2 F_{q,X}(w) = -q^n \|\mathcal{P}\|^2_{L^2(\gamma_d)}\big[k(k-1)S_k(w)D_k(w,v) + k^2 U_k(w,v)^2\big] + o(q^n)D_k(w,v).$$

For some $q_X > 0$, the entropy is strictly concave on the full simplex whenever $0 < q \le q_X$. For $n \ge 4$ the second derivatives have continuous extensions to faces. At $n = 3$ the entropy has a continuous first derivative in squared gains, and the second derivative tends to $-\infty$ along a nonzero line entering a previously zero weight, for sufficiently small q.

By Gaussian moment determinacy on each line, the theorem applies to every fixed non-Gaussian isotropic law with all moments, including log-concave laws. The theorem of Section 4 treats symmetric scalar inputs and two copies. The finite moment extension of Section 17 treats the whole simplex with central symmetry. We use the underlying gain derivatives to allow an odd first discrepancy and arbitrary asymmetry. Luo's counterexamples [7] at each positive signal vary the input with the signal, a distinction quantified in Section 23.4 below.

Proof. Use real gains h_i and residual variance $1 - \sum_i h_i^2$, and write $\mathcal{D}(h)$ for the relative entropy to the standard Gaussian. The output covariance stays I_d. We first give a gain-regularity argument that uses only a moment strictly above the derivative order. On compact gain balls, write the random output as $Z_h = Y_h + \sigma(h)G$, where $Y_h = \sum_i h_i X_i$ and $\sigma(h)^2 = 1 - |h|^2$. For a density p_h of Z_h, define

$$J_i(x) = p_h(x)\mathbb{E}[X_i + \partial_i\sigma(h)G \mid Z_h = x].$$

Differentiating against smooth test functions gives $\partial_i p_h = -\operatorname{div} J_i$. The entropy derivative is $-\int J_i \cdot \nabla \log p_h$. To justify the first derivative and this integration by parts, insert smooth spatial cutoffs. A Gaussian lower envelope bounds the log by $C(1+|x|^2)$; the cutoff error uses at most the third latent moment. The first derivative without the flux formula can also be obtained directly from the Gaussian kernel under this same moment order, as in Section 3.

Differentiate the flux integrand $j-1$ more times for any desired $j \ge 3$. A normalized kernel derivative of order u has latent degree at most u, and a normalized derivative of the flux of order u at most $u+1$. Spatial derivatives add only Gaussian innovation powers. These statements follow by differentiating the Gaussian kernel and its standardized residual, whose gain derivative is linear in the latent inputs and the innovation. Each term in the differentiated entropy integral is the density times a product of such normalized factors of total latent degree at most j. A factor of degree $k > 0$ is bounded in $L^{(j+\epsilon)/k}(p_h)$ by conditional Jensen and independence of the latent inputs and innovation. A degree-zero factor has every finite L^p bound. Hölder has positive exponent left for $1_{\{|Z_h|>R_0\}}$ because $j/(j+\epsilon) < 1$. Uniform tightness on compact gain balls therefore makes the derivative tails vanish. Compact spatial truncation and local kernel differentiation prove C^j entropy regularity. We use $j = 2n$. The argument also applies when the gains and positive Gaussian variance depend smoothly on a single real parameter, with the corresponding velocity in the flux.

The leading density gain jet is $\phi_d \mathcal{P} \sum_i h_i^n$, with every nonconstant coefficient below total degree n zero. This follows by differentiating the characteristic function after its standard Gaussian factor is extracted. Thus

$$\mathcal{D}(h) = \tfrac{1}{2}\|\mathcal{P}\|_2^2 \left(\sum_i h_i^n\right)^2 + R(h),$$

and R has zero gain jet through order $2n$. Only the square of that first density jet enters the entropy coefficient at degree $2n$, since mass kills the linear term. This is a finite Taylor argument. At zero gain all the density jet coefficients are Gaussian times polynomials. Thus the flux derivatives giving

the entropy jet can be integrated by parts at zero, or obtained from spatial cutoffs with Gaussian tail errors, under the same moment assumption.

For any i, fixing the other gains and replacing X_i by a standard normal makes the output independent of h_i. Gaussian matching gives the same density and entropy derivatives through order $n-1$ at $h_i = 0$. Hence $\partial_i^j R = 0$ on that hyperplane for $1 \le j < n$, including their tangential derivatives whenever defined. The leading polynomial has the same vanishing. Use the zero jet at the origin to get $\partial_i^n R = o(|h|^n)$ and $\partial_i^n \partial_j^n R = o(1)$ for $i \neq j$. Integrating in one or two coordinates gives

$$\partial_i R = h_i^{n-1} o(|h|^n), \qquad \partial_{ii} R = h_i^{n-2} o(|h|^n), \qquad \partial_{ij} R = h_i^{n-1} h_j^{n-1} o(1) \quad (i \neq j)$$

in the nonnegative gain orthant, uniformly as $|h| \to 0$.

Substitute $h_i = \sqrt{q w_i}$. The diagonal weight derivative is $q^2(h_i^{-2}\partial_{ii} R - h_i^{-3}\partial_i R)/4$, and the mixed derivative is $q^2 (h_i h_j)^{-1} \partial_{ij} R/4$. The diagonal remainder is $o(q^n) w_i^{k-2}$, and the mixed remainder is $o(q^n) w_i^{k-1} w_j^{k-1}$. Use $(\sum_i |v_i| w_i^{k-1})^2 \le S_k(w) D_k(w, v)$ to sum these bounds. The Hessian of the leading square gives the asserted expansion after negating relative entropy. The lower bound $S_k(w) \ge m^{1-k}$ proves strict curvature along every nonconstant simplex segment in its open part, by restricting to its support face. Entropy is continuous at the endpoints of every such segment, proving strict concavity.

The face factorizations also show the stated regularity. Before estimating at the origin, the integral Taylor formula writes diagonal weight derivatives as h_i^{n-4} times continuous functions of the nonnegative gains, and a first derivative as h_i^{n-2} times a continuous function. The mixed derivatives have continuous factors $h_i^{n-2} h_j^{n-2}$. At $n = 3$, entering a zero weight contributes a term $w_i^{-1/2} v_i^2$ to D_k, so the uniform negative upper bound forces divergence to $-\infty$. □

23.2 The curvature at an activating face

For this section even identical input laws and isotropy are unnecessary. Let the X_j be independent centered vectors with finite $(4+\epsilon)$-th absolute moments for some $\epsilon > 0$, and fix $0 < q < 1$. At a simplex face let w^0 be the base point and v a feasible one-sided tangent. Let $J_0 = \{j : w_j^0 = 0\}$, so $v_j \ge 0$ for $j \in J_0$. Use the path $w(t) = w^0 + tv$. Write p for the density at the face and put

$$B = \sqrt{q} \sum_{j \in J_0} \sqrt{v_j} X_j, \qquad \mathcal{B}_j(p) = \int (\nabla^j p)(x) \log p(x)\, dx \quad (j = 3, 4).$$

A colon denotes contraction of tensors. The reference entropy $F_{\mathrm{G}}(t)$ uses a centered Gaussian with the covariance of B in place of B at the entering slots, keeping the same active inputs and spare Gaussian.

Theorem 23.2. *Writing $F(t)$ for the entropy along this path,*

$$F''(t) = \frac{\kappa_3(B) : \mathcal{B}_3(p)}{8\sqrt{t}} + F_{\mathrm{G}}''(0) - \frac{\kappa_4(B) : \mathcal{B}_4(p)}{12} + o(1) \qquad (t \downarrow 0),$$

where κ_j are cumulant tensors and the reference derivative at zero is one-sided. Moreover

$$\mathcal{B}_3(p) = -\tfrac{1}{2} \mathbb{E}_p[(\nabla \log p)^{\otimes 3}].$$

If the inputs are isotropic and the active weights are all reduced in proportion, $v_j = -\lambda w_j^0$ for $j \notin J_0$, where $\lambda = \sum_{j \in J_0} v_j$, then

$$F_{\mathrm{G}}''(0) = -\frac{\lambda^2}{2} \mathbb{E}_p \|\nabla^2 \log(p/\phi_d)\|_{\mathrm{HS}}^2.$$

Proof. In the variable $s = \sqrt{t}$ the output random vector has the form $A(s^2) + sB + \sqrt{1-q}\, G$, with smooth active gains in $A(s^2)$, and B centered independent of the active inputs. Write $H(s) = F(s^2)$ for $s \geq 0$ and extend via this output to negative s. The flux argument in Section 1 gives C^4 regularity at zero. The first density derivative is zero. The third is $-\kappa_3(B) : \nabla^3 p$: any cross term from an active even gain change would contain one factor of B and its mean vanishes. Consequently $H'(0) = 0$ and $H'''(0) = \kappa_3(B) : \mathcal{B}_3(p)$. At order four the density derivative changes by $\kappa_4(B) : \nabla^4 p$ when passing from the Gaussian-activation reference to the true law. All lower terms that can be paired in the fourth entropy derivative agree: the first derivatives vanish and the second ones agree. Thus $H^{(4)}(0) - H_{\mathrm{G}}^{(4)}(0) = -\kappa_4(B) : \mathcal{B}_4(p)$. These coefficient identities also follow under the stated moment assumption from differentiating the flux. At zero the flux and first density derivative vanish. The remaining term for $H'''(0)$ pairs the second flux derivative with the score. Its cubic part is the contraction of $\kappa_3(B)$ with $\nabla^2 p$, leaving its vector index free. In the comparison of fourth derivatives, all lower flux and density terms agree wherever they enter. The difference of third flux derivatives is the negative contraction of $\kappa_4(B)$ with $\nabla^3 p$, leaving its vector index free. These score pairings give the displayed log integrals by integration by parts. Their tails are controlled by the Gaussian kernel moment bounds. Since H_{G} is even in s, the chain identity

$$F''(s^2) = \frac{sH''(s) - H'(s)}{4s^3} = \frac{H'''(0)}{8s} + \frac{H^{(4)}(0)}{12} + o(1)$$

proves the expansion.

For any direction e, let $\rho_e = \partial_e \log p$. Integration by parts gives $\int \partial_e^3 p \log p = -\mathbb{E}_p \partial_e^2 \rho_e = -\frac{1}{2}\mathbb{E}_p \rho_e^3$. For the last equality use $\mathbb{E}_p \partial_e^2 \rho_e = -\mathbb{E}_p \rho_e \partial_e \rho_e = \frac{1}{2}\mathbb{E}_p \rho_e^3$. The identities extend to tensors by polarization. Polynomial kernel expressions and conditional Jensen give integrable spatial score derivatives, so cutoff integration by parts is valid. In the proportional isotropic case, Gaussian activation is the Ornstein–Uhlenbeck smoothing of p at signal $1 - \lambda t$. Along the OU signal a, differentiation of entropy and relative Fisher information gives $dh/da = -I/(2a)$ and $dI/da = (I + A)/a$, where $A = \mathbb{E}\|\nabla^2 \log(p_a/\phi_d)\|_{\mathrm{HS}}^2$. These are the ordinary OU dissipation identities (Section 15.2). This yields the stated restoring term. □

For two scalar copies with one entering weight the singular coefficient is precisely Luo's $-\mu_3 J_3/16$ when evaluated at the smoothed input [7, Appendix A]. The regular term with symmetric scalar input is the endpoint expression of Abu-Ajamieh [2] and Section A. The theorem allows arbitrary faces and independent vector inputs.

There is also a local statement for the full transverse Hessian. At a fixed boundary point choose one positive pivot weight to absorb the sum constraint. Parametrize the potentially zero slots by their square-root weights s_j and use ordinary weight coordinates for the other nonpivot slots. The entropy $H(s, a)$ is jointly C^4 by the same flux argument, since the pivot stays positive. The trace $H_{s_j} = 0$ at $s_j = 0$ follows from centering and the quadratic variation of the pivot gain. Thus the j-th diagonal transverse Hessian entry is

$$\frac{H_{s_j s_j s_j}(0, s_{-j}, a)}{8 s_j} + O(1)$$

uniformly in a compact neighborhood. The coefficient is evaluated on the nearby face with that one slot removed. Mixed entries are bounded even when the entering weights vanish at different rates. For two entering slots, $H_{s_j s_k}$ vanishes on both coordinate hyperplanes by the same trace identity, so integrating $H_{s_j s_j s_k s_k}$ gives $H_{s_j s_k} = O(s_j s_k)$, which cancels the two square-root chain factors. A mixed entering and ordinary coordinate uses one such integration.

23.3 A sharp third-moment scale for positive singular curvature

For iid isotropic log-concave inputs, the singular term in Theorem 23.2 cannot have positive sign at sufficiently small signal unless the first third cumulant is itself very small. The following estimate gives the scale.

Theorem 23.3. *Fix d, m. There are $q_0, C > 0$ such that for any iid centered isotropic log-concave inputs, positive singular curvature along a feasible activating face direction at $0 < q \le q_0$ requires*

$$\|\kappa_3(X)\| \le Cq^3.$$

Here the norm is the Hilbert–Schmidt tensor norm. The exponent three is sharp already in dimension one and with two copies.

Proof. We show that for any centered isotropic log-concave Y, uniformly over this class in fixed dimension,

$$\mathcal{B}_3(p_{Y_q}) = -q^{3/2}\kappa_3(Y) + O_d(q^{7/2}\|\kappa_3(Y)\| + q^{9/2}). \tag{117}$$

Write $p_{Y_q} = \phi_d(1+u)$. On $|x| \le q^{-\theta}$, with fixed small $\theta > 0$, the Gaussian kernel admits the uniform relative expansion

$$u = q^{3/2}P_3 + q^2P_4 + q^{5/2}P_5 + O(q^3\,\mathrm{poly}(|x|)) \tag{118}$$

through three spatial derivatives. The P_j are the degree-j Gaussian Hermite polynomials determined by the moment discrepancies of Y. For a verification, with gain h the relative Gaussian kernel is

$$(1-h^2)^{-d/2}\mathbb{E}\exp\left(\frac{hx\cdot Y - h^2(|x|^2+|Y|^2)/2}{1-h^2}\right).$$

The centered moment comparison for standardized log-concave inputs gives a common exponential moment in fixed dimension [9]. For $|h| \le \sqrt{q}$ the linear tilt on this ball stays inside that exponential moment range. Every derivative in h through order six and in x through order three is bounded by a fixed polynomial in $|x|$, uniformly. The gain Taylor coefficients are the Gaussian Hermite moments, by the Gaussian generating function or characteristic functions. The mass, mean and covariance give vanishing orders zero through two for u. Choose θ small enough that $|u| \le 1/2$ on this ball, using the polynomial bounds.

Outside this ball the true output and its Gaussian-weighted derivatives contribute less than every power of q to the pairings $\mathcal{B}_3$. In fact the true tail is at most $Ce^{-cq^{-2\theta}} + Ce^{-cq^{-1/2-\theta}}$, by the Gaussian and latent tails. The density derivative divided by the density is a conditional polynomial in the Gaussian innovation, so conditional Jensen, Cauchy–Schwarz and the uniform moments control its product with the quadratic logarithm on this tail. Polynomial Gaussian test pairings have the same property. This justifies using the bulk expansion and extending its polynomial coefficients to the whole Gaussian integral.

For a multi-index ν of degree three, $\int(\partial^\nu p)\log\phi_d = 0$. The linear pairing in $\log(1+u)$ is exactly $-\mathbb{E}\mathrm{He}_\nu(Y_q) = -q^{3/2}\kappa_{3,\nu}(Y)$ up to a superpolynomial tail error. The quadratic part has terms

$$\left\langle \frac{\partial^\nu(\phi_d u)}{\phi_d}, u\right\rangle_{\gamma_d} + \tfrac{1}{2}\langle \mathrm{He}_\nu, u^2\rangle_{\gamma_d},$$

where the pairings at this stage are over the bulk. Cubic errors cost $O(q^{9/2})$ by the polynomial bounds in (118) and the Gaussian moments. The terms from P_3^2, P_4^2, P_3P_5 vanish by parity on the

radial bulk. The terms from P_3P_4 cost $O(q^{7/2}\|\kappa_3(Y)\|)$, since P_3 is linear in $\kappa_3(Y)$. All remaining terms and the remainder in (118) cost $O(q^{9/2})$. This proves (117).

At an iid face the active latent sum is $Y = \sum_{j\notin J_0}\sqrt{w_j^0}X_j$. It is isotropic and log-concave, and $\kappa_3(Y) = c_3\kappa_3(X)$ with $c_3 = \sum_{\text{active}}(w_j^0)^{3/2} \geq m^{-1/2}$. The entering tensor is a positive scalar multiple of $\kappa_3(X)$ when a weight enters. If its pairing in Theorem 23.2 is positive, (117) implies

$$0 < -c_3q^{3/2}\|\kappa_3(X)\|^2 + Cq^{7/2}\|\kappa_3(X)\|^2 + Cq^{9/2}\|\kappa_3(X)\|.$$

Absorb the second term for small q and divide by the nonzero skewness. This yields the asserted q^3 bound. Sharpness follows from the degree-four case in Section 23.4. □

On the line there is a precise limiting constant for a fixed face. Put $c_j = \sum_{\text{active}}(w_i^0)^{j/2}$, and define

$$C_* = \max\{|\kappa_4(X)\kappa_5(X)| :\ X \text{ centered, variance one, log-concave, } \kappa_3(X) = 0\}.$$

The maximum exists by weak compactness, uniform exponential moments and closure of log-concavity. It is positive, for example by the degree-four base laws in Section 23.4. Among laws with a positive singular coefficient at this face, the supremum of $|\kappa_3(X)|/q^3$ tends to

$$\frac{c_4c_5}{2c_3}\,C_*.$$

To see this, in the regime $|\kappa_3(X)| = O(q^3)$ the only contribution of order $q^{9/2}$ beyond the linear term in (117) comes from P_4P_5 in the Hermite multiplication term. The derivative pairing raises degree by three, so the pair of adjacent degrees contributes zero there. Since $\langle \mathrm{He}_3e_4, e_5\rangle = 12\sqrt{5}$ for the normalized Hermites, one obtains uniformly

$$\mathcal{B}_3(p_{Y_q}) = -q^{3/2}c_3\kappa_3(X) + \tfrac{1}{2}q^{9/2}c_4c_5\kappa_4(X)\kappa_5(X) + O(q^5).$$

This and weak compactness give the upper limit. For the reverse direction choose a maximizing zero-skew law. Under exponential tilt, followed by recentering and dividing by the standard deviation, the derivative of the third cumulant at zero tilt is $\kappa_4 \neq 0$. Indeed the third cumulant before standardization has derivative κ_4, and the derivative of the scale factor is multiplied by the zero third cumulant. The moment generating function is finite near zero, so this tilt is available and preserves log-concavity. Prescribe third cumulant aq^3 for any a of the sign of $\kappa_4\kappa_5$ with $0 < |a| < (c_4c_5/(2c_3))C_*$. The fourth and fifth cumulants tend to their maximizing values, and the displayed expansion gives positive singular curvature for all sufficiently small q. This proves the claimed variational value.

Luo's explicit construction already contains a third-moment adjustment of order q^3 before the chosen perturbation amplitude is included [7, Section 3]. Theorem 23.3 gives a class-wide upper restriction, and the next section attains the scale while the laws converge to a fixed non-Gaussian limit.

23.4 A moving endpoint transition at a fixed non-Gaussian limit

Let $n \geq 4$ be even, and write $e_j = \mathrm{He}_j/\sqrt{j!}$ on the line. Choose compactly supported smooth even and odd functions of first Hermite degrees $n, n+1$, respectively, each with positive coefficient in its first degree. Such functions are obtained by cutting off e_n, e_{n+1} and correcting finitely many moments using a compact smooth cutoff times the corresponding even or odd Hermite basis. The

Gram matrices are positive definite on each parity subspace, since a nonzero polynomial does not vanish on an interval. Take a sufficiently small positive linear combination and denote it by g_0, with its normalized coefficients at $n, n+1$ denoted by $a, b > 0$. The density $\phi(1+g_0)$ is strongly log-concave, by smallness of the first two derivatives, and has first discrepancy at degree n.

Choose a compactly supported smooth odd g_3 with zero first Hermite moment and $\langle \mathrm{He}_3, g_3\rangle_\gamma = 1$, by the same finite correction. For each small $q > 0$ take

$$g_q = g_0 + \lambda q^{n-1} g_3, \qquad 0 < \lambda < ab\, 3n\sqrt{n+1}.$$

The density $f_q = \phi(1+g_q)$ is strongly log-concave and converges in every bounded relative C^k norm to the fixed $f_0 = \phi(1+g_0)$. In particular the family stays a positive Hellinger distance from the Gaussian. Write p_q for the OU-smoothed density, and let $u_q = p_q/\phi - 1$. Then $\kappa_3(p_q) = \lambda q^{n+1/2}$.

The third entropy pairing satisfies

$$B_3(p_q) := \int p_q''' \log p_q = q^{n+1/2}(ab\, 3n\sqrt{n+1} - \lambda + o(1)).$$

For verification, expand it to second order in u_q, as in Section 3. Its linear term is $-\langle \mathrm{He}_3, u_q\rangle_\gamma$. In fixed Gaussian Sobolev norms the leading even and odd terms are $aq^{n/2}e_n$ and $bq^{(n+1)/2}e_{n+1}$, with errors $O(q^{(n+2)/2})$ and $O(q^{(n+3)/2})$, respectively. The added degree-three part is $O(q^{n+1/2})$. The derivative pairing of the two leading modes vanishes because the raising operator shifts degree by three. Same-parity pairs vanish. The multiplication term contributes $abq^{n+1/2}\langle \mathrm{He}_3 e_n, e_{n+1}\rangle$, and the latter coefficient is $3n\sqrt{n+1}$ by the Hermite product identity. Higher products are lower order by Gaussian hypercontractivity and the spare smoothing in q. More explicitly, for each fixed derivative order and finite $L^p(\gamma)$ norm, u_q and its derivatives have size $O(q^{n/2})$. Indeed differentiation of the OU image multiplies by $q^{j/2}$ and reduces the minimum Hermite degree by at most j. A fixed-factor extra signal supplies hypercontractivity and any fixed number powers. This also bounds the cubic logarithm remainder by $O(q^{3n/2})$, for example by Hölder with the Gaussian polynomial factors. Since $n \geq 4$, this is lower order.

Let $F_q(t)$ be the two-copy entropy at these input laws and retained signal q. Define the positive constants

$$C_s = \frac{\lambda(ab\, 3n\sqrt{n+1} - \lambda)}{8}, \qquad C_r = a^2\left(\frac{n(n-1)}{2} + 2\,1_{\{n=4\}}\right).$$

Then

$$q^{-n}F_q''(\xi q^{2n+2}) \longrightarrow C_s\,\xi^{-1/2} - C_r \tag{119}$$

uniformly on compact positive intervals of ξ. In fact for small $q > 0$ there is exactly one inflection point near each endpoint and none elsewhere. The one near zero obeys $t_*(q)/q^{2n+2} \to (C_s/C_r)^2$.

To prove the quantitative remainder, write $H_q(s) = F_q(s^2)$, $a_s = \sqrt{1-s^2}$. The relative density ratio under a Gaussian reference, conditional on its normalized sum, is

$$1 + R_s(x) = \mathbb{E}_G[(1 + u_q(a_s x - sG))(1 + u_q(sx + a_s G))].$$

On $|s| \leq 1/4$, its first five s-derivatives and R_s itself have $L^p(\gamma)$ norms $O(q^{n/2})$ for every fixed finite p. Each derivative gives products of derivatives of u_q and fixed-degree polynomials in (x, G). Both Gaussian arguments are standard normal under joint Gaussian measure, so Hölder and conditional Jensen give the assertion from the preceding bounds. The density ratio is bounded away from zero by the small fixed relative sup norm. By mass and variance matching the entropy deficit is the Gaussian average of $(1+R_s)\log(1+R_s) - R_s$. Every differentiated term through order five has at least two factors of size $O(q^{n/2})$: the term with only one derivative factor also has $\log(1+R_s) = O(R_s)$. This proves $\sup_{|s|\leq 1/4} |H_q^{(5)}(s)| = O(q^n)$.

Theorem 23.2, or the same density calculation with U, V distributed as p_q, now gives

$$F_q''(t) = \frac{\kappa_3(p_q)B_3(p_q)}{8\sqrt{t}} - \frac{A(p_q)}{2} - \frac{\kappa_4(p_q)B_4(p_q)}{12} + O(q^n\sqrt{t})$$

for $t \le 1/16$, with $A(p_q) = \int((\log(p_q/\phi))'')^2 p_q$. Here we have used the fifth-derivative estimate in Taylor's formula for $sH_q''(s) - H_q'(s)$. The regular terms satisfy

$$A(p_q) = q^n a^2 n(n-1) + o(q^n), \qquad \kappa_4(p_q)B_4(p_q) = q^n\, 4!a^2 1_{\{n=4\}} + o(q^n).$$

For the first, $(\log(1+u_q))''$ agrees with u_q'' up to $O(q^n)$ in $L^2(\gamma)$ by the L^4 bounds and positivity of the relative ratio. The density factor costs only a lower order term by Hölder. For the second, the fourth cumulant is zero for $n > 4$; for $n = 4$, the linear term in B_4 equals $\kappa_4(p_q)$, with quadratic error $O(q^n)$. This proves (119).

To see uniqueness, the same gain expansion gives

$$F_q'''(t) = -\frac{\kappa_3(p_q)B_3(p_q)}{16t^{3/2}} + O(q^n t^{-1/2}).$$

This follows by differentiating the exact chain formula before taking Taylor expansions through order four in H_q', H_q'', and H_q'''. Thus F_q'' is strictly decreasing on a fixed compact crossover interval containing $(C_s/C_r)^2$, for small q. At smaller scaled weights its singular term dominates the regular terms. At larger scaled weights, up to a fixed small t_0, the regular negative term dominates by the uniform $O(q^n\sqrt{t})$ bound. On every fixed compact interval in $(0,1)$, the density ratio has the following leading term in the first two weight derivatives as well: $1 + aq^{n/2}(t^{n/2} + (1-t)^{n/2})e_n$. This follows from the linear OU projection and the $O(q^n)$ product term, using the same conditional argument. Expanding entropy gives the strictly negative bulk curvature limit from the square of that linear term. Reflection of the weights gives the same conclusions near one.

23.5 Unequal symmetric inputs

For distinct input laws, even symmetry of both laws does not ensure eventual concavity. In fact there are fixed smooth strongly log-concave symmetric variance-one laws A, B, arbitrarily close to Gaussian in bounded relative C^k norms, with positive endpoint curvature at every $0 < q < 1$ when the active law is A and the entering law is B.

To prove this, choose a smooth compactly supported even g of minimum Hermite degree four with $k_4 = \langle \mathrm{He}_4, g\rangle_\gamma > 0$. Take relative densities $1 + a_0 g, 1 - b_0 g$ with $0 < a_0 \ll b_0 \ll 1$. They are positive strongly log-concave and match mass and variance. For $u_q = T_{\sqrt{q}}g$, the Hermite degree bound and Gaussian derivative identities give

$$\|u_q\|_2 + \|u_q'\|_2 + \|u_q''\|_2 + \|(\phi u_q)^{(4)}/\phi\|_2 \le Cq^2 \qquad (0 < q \le 1).$$

The fixed finite Gaussian Sobolev norms at signal one pay for the derivatives. Ordinary OU contraction also gives bounded spatial derivatives in sup norm. For the active smoothed density $p = \phi(1 + a_0 u_q)$ these estimates imply

$$A(p) \le Ca_0^2 q^4, \qquad B_4(p) = a_0 q^2 k_4 + O(a_0^2 q^2) \ge a_0 q^2 k_4/2$$

uniformly for small a_0. For the second estimate expand the logarithm in the defining pairing; the linear term is $\langle \mathrm{He}_4, a_0 u_q\rangle_\gamma$, and Cauchy–Schwarz and the uniform relative bound give the stated

error. The entering smoothed law has fourth cumulant $-b_0q^2k_4$, while both third cumulants vanish by symmetry. Theorem 23.2 therefore bounds the finite endpoint curvature below by

$$q^4(a_0b_0k_4^2/24 - Ca_0^2) > 0$$

at fixed sufficiently large ratio b_0/a_0. Sending both amplitudes to zero at that ratio gives the claimed Gaussian-relative closeness.

24 The extremal fifth moment of skew-zero log-concave laws

24.1 Result

Let $\mathcal{L}$ consist of real log-concave laws with mean zero, variance one and third moment zero. Write $m_j(X) = \mathbb{E}X^j$. For $t > 0$ and $a > 0$, let $Y_{a,t}$ have normalized density proportional to

$$e^{ax}1_{\{-t/a \le x \le 0\}} + e^{-x}1_{\{x>0\}}.$$

There is a unique $a = a(t)$ for which $Y_{a,t}$ has zero third central moment. Let X_t be $Y_{a(t),t}$ centered and divided by its standard deviation. Define $K(t) = m_4(X_t)$, $S(t) = m_5(X_t)$.

Theorem 24.1. *The fourth-moment range in $\mathcal{L}$ is $[9/5, 6]$. The two ends are attained only by the variance-one uniform and Laplace laws, respectively. The map $t \mapsto K(t)$ is a strictly increasing bijection from $(0, \infty)$ onto $(9/5, 6)$. For any $X \in \mathcal{L}$ with $m_4(X) = K(t)$,*

$$|m_5(X)| \le S(t).$$

Here $S(t) > 0$, and equality holds only for X_t and its reflection.

The scalar constant obtained from the endpoint expansion in Section 23 is therefore

$$\max_{X \in \mathcal{L}} |\kappa_4(X)\kappa_5(X)| = \max_{t>0} |K(t) - 3|\, S(t). \tag{120}$$

The moments on the right are explicit in t and one uniquely specified positive root. Both endpoint limits of the objective are zero, so a maximizer occurs at a finite positive t. The comparison concerns nonsymmetric laws with zero third moment. The moment problem of Eskenazis, Nayar and Tkocz [18] concerns symmetric log-concave inputs, for which fifth moments vanish. Generalized localization [19] gives a broader framework for moment constraints. The specific one-sided two-piece form here follows from a sign comparison that also proves uniqueness at each fourth moment.

24.2 Moment comparison by signs

We record the elementary sign-change argument. Suppose two densities f, g have identical moments through degree $k-1$. If their difference is nonzero and has at most k sign changes, it has exactly k. In fact, with fewer changes, choose one point at each change and form a polynomial whose roots are those points. Orient its sign so its product with $f - g$ is nonnegative and nonzero on a set of positive measure. Its integral cannot vanish, contradicting equality of the lower moments. With exactly k changes the same argument with a monic polynomial of degree k determines the sign of $\int x^k(f - g)$ by the last sign of $f - g$. Zeros and intervals of equality may be skipped. In all comparisons below the densities are absolutely continuous with finite moments.

On every interval where g is positive and log-affine, the positive set of $f - g$ is an interval, because f is log-concave. For the centered variance-one Laplace law g, there are at most two

positive intervals, one on each side of its knot. Equal moments through degree three therefore force exactly four sign changes for $f \neq g$, with last sign negative. Thus $m_4(f) < m_4(g) = 6$. For the variance-one uniform law there is one possible positive interval inside its support, and the exterior differences are nonnegative. If $f \neq g$, equality through degree three forces exactly four changes, with last sign positive. Hence $m_4(f) > m_4(g) = 9/5$. In particular every $X \in \mathcal{L}$ other than the uniform law gives positive mass strictly to each side of the interval $[-\sqrt{3}, \sqrt{3}]$: without one of those positive exterior parts the required four sign changes are impossible.

For the standardized one-sided two-piece density g from Section 1, there can be two positive intervals on its support, together with a possible positive part to the left of its finite endpoint. If the exterior part is absent, there are at most four changes. Otherwise the first sign is positive and there are at most five changes. Thus equality of moments through degree four for a different log-concave density f forces five changes, with last sign negative. This proves $m_5(f) < m_5(g)$, with strict inequality whenever $f \neq g$. Comparison with the reflection of g, which has the same first four moments at zero skew, shows $m_5(g) > 0$. Reflection of the comparison gives both signs of the claimed fifth-moment bound.

24.3 The skew-zero curve

For $j \geq 1$ put

$$G_j(t) = \int_0^t \frac{x^{j-1}e^{-x}}{(j-1)!}\,dx = 1 - e^{-t}\sum_{i=0}^{j-1}\frac{t^i}{i!}.$$

The unnormalized moments of $Y_{a,t}$ are

$$M_j = j!\left(1 + \frac{(-1)^j G_{j+1}(t)}{a^{j+1}}\right), \qquad m_j(Y_{a,t}) = M_j/M_0.$$

Set $A = G_1(t), B = G_2(t), C = G_3(t), D = G_4(t)$. The third central moment is zero precisely when $M_3M_0^2 - 3M_2M_1M_0 + 2M_1^3 = 0$. After multiplying by $a^6/2$, this becomes

$$\begin{aligned} P_t(a) = {} & a^6 + 3Aa^5 + 3A^2a^4 + 3(AB - C)a^3 \\ & + 3(B^2 - AC - D)a^2 + 3(BC - 2AD)a - 3A^2D + 3ABC - B^3. \end{aligned} \tag{121}$$

To verify its coefficient signs, first $AB > C$: the sum of independent gamma variables of shapes one and two and rate one has shape three, and the event that their sum is at most t is strictly contained, up to positive probability, in the event that each is at most t. Next let T be a rate-one exponential restricted to $[0, t]$ and normalized. Any nonincreasing density on a positive ray is a mixture of uniforms on intervals $[0, s]$: if f is the density, take the mixing measure $-s\,df(s)$ including downward jumps, and integrate $1_{[0,s]}/s$. Thus one can write $T = UZ$ with independent uniform U on $[0, 1]$ and nonnegative Z. For this nonconstant density Z is nondegenerate. In particular

$$\mathbb{E}(T - \mathbb{E}T)^3 = \tfrac{1}{4}\mathbb{E}[Z(Z - \mathbb{E}Z)^2] > 0, \qquad 2\mathbb{E}T^3 > 3\,\mathbb{E}T\,\mathbb{E}T^2.$$

The second inequality is just $\mathbb{E}Z^3 > \mathbb{E}Z\,\mathbb{E}Z^2$, obtained by the positive covariance of the increasing powers Z, Z^2. Translating via $(\mathbb{E}T, \mathbb{E}T^2, \mathbb{E}T^3) = (B/A, 2C/A, 6D/A)$, the coefficient of a in (121) is negative, and the constant is $-A^3\mathbb{E}(T - \mathbb{E}T)^3/2 < 0$. The coefficient of a^2 can have either sign. Consequently the coefficient sequence from high degree to low has exactly one sign variation. By Descartes' rule and the signs at zero and infinity, there is exactly one positive root, and it is simple. The root depends continuously, indeed smoothly, on $t > 0$.

As $t \to \infty$, all $G_j(t) \to 1$. The limit polynomial has a unique positive root equal to one, as follows either by direct substitution and the same sign count, or from the zero skewness of the two-sided exponential law with equal rates. Bracketing that simple root yields $a(t) \to 1$. The standardized law tends with its moments to the variance-one Laplace law.

As $t \downarrow 0$, all lower coefficients in (121) vanish, so $a(t) \to 0$ by root bounds for a monic polynomial. Let $L = t/a(t)$. If some subsequence of L were bounded, the raw laws would converge with their moments to a density constant on $[-L_0, 0]$ and exponential on the positive ray, with $L_0 \geq 0$. This law, translated to its finite left endpoint, is nonincreasing and not uniform. The same uniform-mixture argument above makes its third central moment strictly positive, contradicting zero skewness. Thus $L \to \infty$. On dividing the raw variable by L, its left density is proportional to e^{ty} on $[-1, 0]$, and the positive tail has mass $O(1/L)$ with exponential rate L. The rescaled law and its moments tend to uniform on $[-1, 0]$, giving the standardized uniform limit.

The standardized laws X_t are distinct: the product of the positive left log slope and the length of that piece is the invariant parameter t. If two distinct parameters gave the same fourth moment, the strict fifth-moment comparison in Section 2 applied in both directions would be a contradiction. Thus $K(t)$ is continuous and injective. Its endpoint limits make it strictly increasing onto $(9/5, 6)$. This completes Theorem 24.1 and (120).

The slope $a(t)$ is itself strictly increasing. To verify this, keep a fixed and extend the left support by increasing $L = t/a$. At a zero-skew law, the derivative of its third central moment in L is a positive multiple of

$$(-L - \mu)((-L - \mu)^2 - 3\sigma^2),$$

where μ, σ^2 are the raw mean and variance. This follows by differentiating expectations as a boundary mass is added, including the derivative of the mean. The standardized left endpoint is strictly below $-\sqrt{3}$ by the uniform comparison, so this derivative is negative. The partial derivative in a at fixed t is positive at the root, since $P_t(a)$ crosses zero with a simple positive slope and the skew numerator differs by a positive factor. Implicit differentiation gives $a'(t) > 0$.

24.4 A certified estimate of the constant

The compact numerical bound uses only integer interval arithmetic. We first control the parameters outside its range.

For $0 < t \leq 1/40$, put $u = a/t$ and translate and scale to $Z = Y_{a,t}/(t/a) + 1/2$. Its unnormalized left density on $[-1/2, 1/2]$ is $e^{-t(1/2-z)}$. The right part is a mass u times $1/2 + uE$ with E standard exponential. The left mass is in $[1 - t/2, 1]$, the even left moments are bounded by the uniform moments, and an odd left moment of order j lies between zero and $t \int_{-1/2}^{1/2} z^{j+1}\, dz$. For the last bound pair z with $-z$ and use that the exponential is 1-Lipschitz on the nonpositive axis.

At $u = .05$ the unnormalized mass and first three moments obey

$$1.0375 \leq N \leq 1.05, \qquad .0275 \leq m_1 \leq .0275 + t/12, \qquad m_2 \leq 1/12 + .01525, \qquad m_3 \geq .0085375.$$

These bounds follow by adding $u\mathbb{E}(1/2 + uE)^j$ to the left moments. They give the positive centered third numerator

$$\begin{aligned} m_3 N^2 - 3m_2 m_1 N + 2m_1^3 &\geq .0085375(1.0375)^2 \\ &\quad - 3(1/12 + .01525)(.0275 + .025/12)1.05 + 2(.0275)^3 \\ &> 4 \cdot 10^{-5}. \end{aligned}$$

By the sign of the sextic across its unique positive root, the actual u is below .05. For such u, using the left mass lower bound .9875 and the same moment estimates gives

$$0 \le \mu = \mathbb{E}Z < .031, \quad .0773 < \mathbb{E}Z^2 < .1, \quad 0 \le \mathbb{E}Z^3 < .009, \quad \mathbb{E}Z^4 < .018, \quad 0 \le \mathbb{E}Z^5 < .003.$$

For example, the fourth and fifth right unnormalized moments are at most .0048325 and .002770625, while the left ones are at most $1/80$ and $t/448$. The lower second-moment bound uses $(1-.025)/(12\cdot 1.05) > .0773$. Thus $\operatorname{Var} Z > .076$. Binomial centering gives a fourth central moment below $.018 + 6(.031)^2(.1) < .019$. The fifth central moment is at most $.003 + 10(.031)^2(.009) + 4(.031)^5 < .0031$ and at least $-5(.031)(.018) - 10(.031)^3(.1) > -.0029$. Both standardized cumulants have absolute value below 2 (use the fourth-moment lower bound 1), so the objective is below 4.

For $t \ge 20$, first take $a \in [.999, 1.001]$. The untruncated law is $E_1 - E_2/a$ with independent standard exponentials. Its raw moments are

$$r_j^\infty = j!(1 + (-1)^j a^{-j-1})/(1 + 1/a).$$

The law with finite left endpoint is the conditional law after deleting $x < -t/a$. For $1 \le j \le 5$ its deleted absolute raw moment is bounded by

$$\frac{a^{-j-1}}{1 + 1/a} \frac{e^{-t} t^j}{1 - j/t}.$$

This follows from the incomplete gamma integral, using $(1 + y/t)^j \le e^{jy/t}$ after shifting its lower limit to zero. Including the normalization cost bounds the raw moment difference by $2e^{-t}t^j$. Since $t^j e^{-t}$ decreases and $e^{-20} < 3 \cdot 10^{-9}$, the five differences are bounded by

$$1.2 \cdot 10^{-7}, \quad 2.4 \cdot 10^{-6}, \quad .000048, \quad .00096, \quad .0192.$$

The explicit untruncated moments give $|r_1^\infty| < .0011$, $1.98 < r_2^\infty < 2.02$, $|r_3^\infty| < .014$, $r_4^\infty < 24.3$, and $|r_5^\infty| < .4$. The centered third moment changes by less than .00006, by the binomial formula. The untruncated third cumulant $2 - 2/a^3$ has opposite signs of magnitudes $> .0059$ at the two ends of the bracket. Thus the root $a(t)$ is in this bracket. The resulting variance is above 1.97, the central fourth moment is below 24.4, and the central fifth moment is below .57 in absolute value by the same raw bounds and binomial centering. The standardized fourth and fifth cumulants then have absolute values below 4 and .11, respectively, giving an objective below .44 on this range.

24.4.1 Compact interval verification

For compact computation it is convenient to scale the raw support to $[-1, \infty)$, writing $u = a/t$. Its unnormalized raw moments are

$$M_j^* = (-1)^j \int_0^1 x^j e^{-tx}\, dx + j! u^{j+1}.$$

The left integrals are decreasing functions of t. On any closed parameter interval the monotonic slope $a(t)$ permits enclosing it between the two endpoint root brackets. Each endpoint bracket is certified by the sign of the third central moment numerator, with the unique root proved above. Given intervals for the raw moments, let $V = M_2^* M_0^* - (M_1^*)^2$. At zero skewness the fourth excess cumulant and the fifth cumulant have respective numerators

$$N_4 = M_4^*(M_0^*)^3 - 3(M_2^*)^2(M_0^*)^2 + 2(M_1^*)^4,$$
$$N_5 = M_5^*(M_0^*)^4 - 5M_4^* M_1^*(M_0^*)^3 + 20M_2^*(M_1^*)^3 M_0^* - 16(M_1^*)^5.$$

Their product in absolute value is $|N_4N_5|/V^{9/2}$. The formulas follow by binomial centering and eliminating $m_3 = 3m_1m_2 - 2m_1^3$ for normalized raw moments. The supplementary replay subdivides $[.025, 20]$ until this expression has the desired interval upper bound on every slab. For e^{-t} at a rational endpoint it inverts a positive Taylor sum for e^t with an explicit geometric tail upper bound. The left moments then follow by integration by parts. All interval endpoints are rounded outward on an integer grid, and square roots are enclosed using integer square roots. Floating-point root proposals serve only to find brackets whose signs are then verified with the intervals.

The replay certifies a strict lower witness at the rational parameter $t = 5.1008$. Using the same interval method, it excludes all t outside $[4.9, 5.3]$ by bounding their objective below the witness. On $[4.9, 5.3]$ the signed product is positive and its second derivative is strictly negative, as checked by interval differentiation along the implicit zero-skew curve. The first derivative is positive at 5.100823412 and negative at 5.100823413. Consequently there is exactly one maximizing parameter, and a rigorous enclosure is

$$5.100823412 < t_* < 5.100823413, \qquad 6.1629504185 < C_* < 6.1629504186.$$

The product at the left endpoint of this bracket gives the refined lower bound. Its tangent line there bounds the concave product from above throughout the bracket. The derivative intervals use the implicit formula $a' = -\partial_t P_t(a)/\partial_a P_t(a)$ and its differentiated version, or equivalently the equation for the raw skew numerator in the scaled variable $u = a/t$. The complete certificate can be replayed with

```
python interval_replay.py -global-bound -unique -cap 6.17.
```

Thus the sharp constant in the scalar entropy application is determined by (120) and enclosed at the displayed precision.

25 The signed moment problem for log-concave laws

25.1 Statement of the results

Let $\mathcal{L}$ be the set of log-concave probability laws on $\mathbb{R}$ with mean zero and variance one. Write $m_j(f) = \int x^j f(x)\,dx$ for a density f of such a law, and let

$$K_j = \{(m_3(f), \dots, m_j(f)) : f \in \mathcal{L}\} \subset \mathbb{R}^{j-2}, \qquad K_2 = \{()\}.$$

For a prescribed $u \in K_{n-1}$, denote the fiber by $\mathcal{L}(u)$. A piecewise log-affine density is positive on an open interval and has a continuous, piecewise affine log density there, with finitely many maximal affine pieces. Its values at the endpoints are immaterial. Let p be the number of pieces and e the number of finite support endpoints.

Theorem 25.1. *Let $n \geq 3$ and $u \in K_{n-1}$. The fiber $\mathcal{L}(u)$ is a singleton if and only if it contains a piecewise log-affine density with $2p + e < n$. This occurs exactly when u is on the boundary of K_{n-1} in $\mathbb{R}^{n-3}$. Otherwise there are unique densities f_+ and f_- attaining the maximum and minimum of m_n. Their support and number of pieces have the following form.*

	f_+	f_-
$n = 2k$	k *pieces on* $\mathbb{R}$	$k-1$ *pieces on a bounded interval*
$n = 2k+1$	k *pieces on* (A, ∞)	k *pieces on* $(-\infty, B)$

Each specified piecewise log-affine form identifies a unique density in the fiber. The same two laws uniquely maximize and minimize $\int \psi f$ for every $\psi \in C^n(\mathbb{R})$ with $\psi^{(n)} > 0$ and $\int |\psi| f_\pm < \infty$, among laws in the fiber for which the integral is absolutely convergent.

Theorem 25.2. *For each $n \geq 3$, all intermediate values between the minimum and maximum of m_n in a fiber occur. These minimum and maximum values are continuous functions on K_{n-1} and are real analytic on its interior. As a result, K_n is homeomorphic to the closed unit ball of dimension $n-2$.*

Eskenazis, Nayar and Tkocz [18] solve the corresponding extremization and moment body problems for symmetric log-concave laws with prescribed absolute moments at real exponents. In the signed problem on the whole line, the position of the finite endpoints determines the direction of the inequality. Curvature perturbations give the order bound $2p + e \leq n$ for every local extremum. Generalized localization under multiple constraints is developed by Fradelizi and Guédon [19]. Chen et al. [20] study discrete analogues with piecewise geometric extremizers. Eitan [21, Theorem 13] proves the third-derivative comparison with shifted exponentials. His skew-matching step in (2.13), carried over to the log-concave case, also gives the order-four upper comparison at fixed skewness.

25.2 Compactness and sign comparisons

The class $\mathcal{L}$ is weakly compact, and all polynomial moments are continuous on it. For completeness, denote the survival probability by S_f. It is log-concave by integration of a log-concave function over a half-line. Chebyshev's inequality gives $S_f(-2) \geq 3/4$ and $S_f(2) \leq 1/4$. The secant estimate for $\log S_f$ gives $S_f(2+t) \leq \frac{1}{4}3^{-t/4}$ for $t \geq 0$. Reflection gives the other tail. Tightness follows, and weak limits retain their first two moments by these bounds. Log-concavity is preserved under weak limits, and a one-dimensional log-concave law of positive variance has a log-concave density. These standard facts, including log-concavity of the integrated tail, are reviewed in [10]. In particular, all fibers are compact.

Lemma 25.3. *If g is piecewise log-affine with p pieces and e finite endpoints, then for any log-concave density f the difference $f - g$ has at most $2p + e$ sign changes. At equality, its last sign is positive if the right endpoint of the support of g is finite, and negative otherwise. If $f \neq g$ and the first n signed moments of f and g (orders $0, \dots, n-1$) agree, then $f - g$ has at least n sign changes whenever it has finitely many.*

Proof. Sign changes are counted after discarding zero sets, with densities specified up to null sets. On each affine piece of $\log g$, the set where $f - g > 0$ is an interval, since $\log f - \log g$ is concave there with $\log 0 = -\infty$. Each such interval contributes at most two changes in the sign sequence. Each allowed exterior half-line has only a positive sign if it has any sign at all, and contributes at most one change. At equality every interior positive interval contributes both changes and every allowed exterior half-line contributes one. This gives the stated last sign.

If there are $r < n$ changes, choose ordered real cuts $s_1, \dots, s_r$ separating consecutive opposite signs. The product $\prod_i (x - s_i)$, times the last sign of $f - g$, has the same sign as $f - g$ wherever the difference is nonzero, up to a null set. The integral of their product is strictly positive. Moment agreement makes the integral zero, a contradiction. □

If $g \in \mathcal{L}(u)$ and $2p + e < n$, the lemma shows the fiber is a singleton. If $2p + e = n$ and $f \in \mathcal{L}(u)$ differs from g, let $s_1, \dots, s_n$ be separating cuts. For a test ψ as in Theorem 25.1, its interpolant of degree at most $n - 1$ at these cuts has remainder with the sign of $\prod_i (x - s_i)$ at each other x, by the interpolation remainder formula. Multiplying by $f - g$ and integrating gives strict inequality in the direction of the last sign of $f - g$. Thus an infinite right endpoint for g identifies a unique maximum, and a finite right endpoint identifies a unique minimum. This also applies to $\psi(x) = x^n$.

25.3 Curvature of a locally extremal density

We first allow the density to be unnormalized, imposing the constraints $m_j = u_j$ for $0 \le j < n$, with $u_0 = 1, u_1 = 0, u_2 = 1$. Local extrema refer to the weak topology on the constrained densities. Every finite-dimensional variation used below is continuous in that topology.

Lemma 25.4. *Any local extremum of m_n in a fiber has a piecewise log-affine density with $2p+e \le n$.*

Proof. Write $f = e^{-V}$ on the interior (A, B) of its support. The convex function V has a nonnegative locally finite distributional second derivative ν on this interval. In addition to varying V by constants and linear terms, allow the following changes of curvature on compact subsets of (A, B). For any finitely many continuous functions ψ_i, ξ_j compactly supported in (A, B), take a compact interval J containing their supports in its interior. Replace $\nu|_J$ by the pushforward of $(1 + \sum_i \eta_i\psi_i)\nu|_J$ under $t \mapsto t + \sum_j \delta_j\xi_j(t)$. Take both signs of sufficiently small parameters. The new curvature is nonnegative and lies in (A, B). The corresponding change in V is the integral of $(x-t)_+$ against the change in curvature. The change is bounded by a constant times the parameter size times $1 + |x|$, so the exponential tails give integrable domination of all moment derivatives. For a fixed t, the hinge is differentiable in the parameters for Lebesgue almost every x, and its difference quotients are bounded by the corresponding displacement quotients. Fubini's theorem and dominated convergence therefore give continuously differentiable moment maps. This holds also when ν has atoms.

For each such finite family the equality constrained multiplier rule gives a nonzero vector of coefficients of $m_0, \dots, m_n$ whose combined first variation vanishes. Indeed, if the constraints have full derivative rank then this follows from the implicit function theorem. Otherwise a nontrivial dependence of their derivatives suffices. Normalize the vector to have Euclidean norm one. For each finite family the allowed vectors are a closed subset of the unit sphere. They have the finite intersection property because the union of finitely many families is also an allowed family. There is thus a single nonzero polynomial P of degree at most n annihilating all the moment variations. In particular,

$$\int Pf = \int xPf = 0.$$

Define

$$H(t) = \int_t^B (x-t)P(x)f(x)\,dx, \qquad H'(t) = -\int_t^B P(x)f(x)\,dx.$$

The variations of curvature mass and its position give $\int H\psi_i\,d\nu = 0$ and $\int H'\xi_j\,d\nu = 0$. Both continuous functions vanish on the support of ν. They also tend to zero at either support endpoint. At the left endpoint use the two affine equations to write each as a left integral, and use exponential tails at infinite ends.

Choose N distinct ordered points in the support of ν. In each of the $N+1$ open intervals between those points and the endpoints, the vanishing of H, H' at both ends implies $\int Pf = \int xPf = 0$ within the interval. For instance, at finite cuts use the definition of H and the known integrals beyond the right cut. For intervals touching an infinite endpoint use the corresponding one-sided formulas for H, H'. Since $f > 0$ inside, the polynomial P must have at least two sign changes in each such interval. Otherwise a polynomial of degree at most one with its sign would contradict these two equations. This bounds N by the degree of P, so ν has finite support and V is piecewise affine.

With finitely many knots $t_1 < \cdots < t_{p-1}$, write

$$V(x) = c + ax + \sum_{i=1}^{p-1} b_i(x - t_i)_+, \qquad b_i > 0.$$

Use as parameters c, a, b_i, t_i, and each finite endpoint. When an endpoint moves, restrict or extend the adjacent affine formula for V. Parameters range over an open set: knots stay ordered and inside the support, jumps stay positive, and the slope on an infinite end stays strictly integrable. All moments are smooth in this parametrization. Reapply the finite multiplier rule. The derivatives in c, a, b_i, t_i give the same affine, $H(t_i)$, and $H'(t_i)$ equations for the new polynomial P. If A is finite its derivative gives $-P(A)e^{-V(A+)} = 0$, and if B is finite its derivative gives $P(B)e^{-V(B-)} = 0$. The endpoint values of the density are positive and finite from within. There are at least two roots of P in each open affine piece and one at each finite endpoint. This proves $2p + e \leq n$. □

25.4 Perturbations preserving the lower moments

Lemma 25.5. *At a piecewise log-affine density of complexity $2p+e = n$, the derivative of the map from the n shape parameters above to $(m_0, \ldots, m_{n-1})$ is invertible.*

Proof. If a nonzero linear combination of the derivative rows vanished, the associated polynomial of degree at most $n - 1$ would satisfy precisely the affine, knot and finite endpoint equations from the preceding proof. It would therefore have at least $2p + e = n$ distinct roots. □

At any such density we may add a hinge of small positive curvature at a fixed point inside one of its affine pieces. The implicit function theorem, with the original n shape parameters as unknowns, permits this variation while leaving the first n moments (orders 0 through $n - 1$) fixed. For the differentiable equations the new hinge coefficient can have either sign. Taking it positive keeps the density log-concave. The existing jumps and the integrability of the tails persist by strict inequalities. The density changes because it acquires a new knot. In particular, a density of complexity n lies in a nonsingleton fiber.

Compactness gives a global maximum and minimum in every fiber. Lemma 25.4 and the sign comparison characterize singleton fibers and force complexity n at all extrema in a nonsingleton fiber. The global maximum and minimum have different values: otherwise both sign comparisons at the maximum and another law in its fiber would give a strict inequality. The global maximum must therefore have an infinite right endpoint, and the minimum a finite right endpoint. The parity of $n = 2p + e$ determines the other endpoint. The sign comparison yields uniqueness and the comparisons for all the claimed tests.

When the fiber is nonsingleton, the invertible moment derivative at either extremum gives an open neighborhood of attainable moment values. Fixing the mass, mean and variance gives a neighborhood of u in $\mathbb{R}^{n-3}$. If the fiber is a singleton, its density has complexity $c_0 < n$. We have $c_0 \geq 3$ since a single affine potential on the whole line cannot give an integrable density. The comparison at order c_0, with the lower moments fixed, places m_{c_0} at one end of its range. Changing that coordinate in the forbidden direction gives infeasible vectors arbitrarily close to u. The compactness of the feasible set gives the boundary assertion of Theorem 25.1.

25.5 Intermediate moments and topology

Consider any local extremum of m_n in a nonsingleton fiber. Its complexity equals n by Lemma 25.4 and the singleton comparison. Lemma 25.5 allows a unique monic polynomial

$$P(x) = x^n - \sum_{j=0}^{n-1} \lambda_j x^j$$

for which the moment combination in P has zero derivative in each original shape parameter. By the root count, it has exactly two roots in each open affine piece, and a root at each finite endpoint,

all simple. Fix a point t in the last affine piece after its last interior root. Add the hinge $\epsilon(x-t)_+$ and adjust the original shape parameters to keep the first n moments fixed as above. The right derivative of the resulting m_n at zero is

$$-\int_t^B (x-t)P(x)f(x)\,dx.$$

The derivatives from the original shape parameters vanish in the polynomial combination, while the lower moments are constant, which gives this formula. If $B=\infty$ then $P>0$ to the right of t and the derivative is strictly negative, excluding a local minimum. If $B<\infty$, then $P<0$ between t and the largest root B, and the derivative is strictly positive, excluding a local maximum. Hence any local extremum has the support pattern appropriate to its direction and is global by the sign comparison.

The continuous image of a fiber under m_n is compact. If this subset of $\mathbb{R}$ has a gap between attainable values, take a law attaining the lower endpoint of the gap. It is a local maximum, because in a sufficiently small weak neighborhood its moment can vary only within less than the gap length. There is a law with strictly larger moment, a contradiction. Thus the range is a closed interval, as is also true for a singleton fiber.

Denote the endpoint values by $a(u)$ and $b(u)$. At an interior point of K_{n-1}, use the shape parameter charts for each extremizer and Lemma 25.5. The moment map is real analytic. Indeed, each finite piece is an integral of a polynomial times an exponential of an affine function, with analytic endpoints and coefficients. A linear change to a fixed finite integration interval proves analyticity even if the slope is zero. Each infinite piece has a nonzero slope of the appropriate sign. Its integral is a finite sum of analytic terms obtained by integrating a polynomial against an exponential on a half-line. The analytic inverse function theorem gives locally analytic extremizing parameters, as the sign comparison identifies the perturbed laws as the extremizers. At a boundary point, compactness and uniqueness of the law in that fiber ensure that either family of extremizing laws converges weakly to this law as u tends to the point. Moment continuity gives continuity of the endpoint values there.

Start from the singleton K_2, the zero-dimensional ball. If h is a homeomorphism from the closed unit ball of dimension $n-3$ to K_{n-1}, choose it to carry its interior onto the interior of K_{n-1}. Over each interior base point x, send the interval

$$[-\sqrt{1-|x|^2},\sqrt{1-|x|^2}]$$

affinely onto $[a(h(x)),b(h(x))]$, with first coordinate $h(x)$. At a boundary base point both intervals collapse, and continuity there follows from continuity of the two endpoint values. The resulting map is a continuous bijection between compact and Hausdorff spaces, hence a homeomorphism. Its interior goes to the points over the interior of K_{n-1} with strict inequalities in the last coordinate. Those are precisely the interior points of K_n, since the endpoint functions are continuous and any open set in K_n projects to an open set in the base. Induction proves Theorem 25.2.

25.6 Gaussian moment matching

Corollary 25.6. *Let $n\geq 3$ and let G be standard Gaussian. Among real log-concave laws matching the moments of G through order $n-1$, the nth cumulant has a closed interval of possible values containing zero in its interior. The two endpoints are attained by the unique laws in the support patterns of Theorem 25.1. For even n, both endpoint laws are symmetric. For odd n, they are reflections of each other.*

Proof. The Gaussian belongs to the fiber and has no affine pieces in its log density. Hence the fiber is nonsingleton and the Gaussian differs from both extremizers. The strict comparison puts its moment in the interior of the range. The difference $m_n(f) - m_n(G)$ equals the nth cumulant of f, since the lower moments agree and Gaussian cumulants of order at least three vanish. The range assertion follows from Theorem 25.2. Reflection preserves this fiber. It preserves the objective when n is even and reverses it when n is odd, so uniqueness gives the symmetry statements. □

Corollary 25.6 concerns the moment-matching conditions used in the Gaussian smoothing and entropy curvature problem of Sections 2–7. It gives both extremal values of the first cumulant left free by those conditions through finite-dimensional moment equations. The conclusion is about the full log-concave class. The bounded perturbations of Gaussian densities used to construct the entropy obstructions in Sections 7 and 17 impose additional regularity and size conditions. At order five, prescribing the fourth moment as well as zero skew gives the specific two-piece family analyzed explicitly in Section 24.

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

$$0 \le g \le M := \frac{1}{1-q}. \tag{122}$$

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

$$F_{tt}(q,0) = -\frac{A}{2} - \frac{\kappa_4(U)}{12}B, \qquad B = \mathbb{E}g^2 - \frac{1}{3}\mathbb{E}\rho^4. \tag{123}$$

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 (123). 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

$$K \le 1 + (1+2C)A \le 116/95 < 5/4. \tag{124}$$

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 (124). Combining the two bounds gives $|\kappa_4(U)B| < (189/32)A$. Equation (123) 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

$$C_{ij} = \operatorname{Cov}_\gamma(G^{2i}, \psi_j(G)), \qquad 1 \leq i, j \leq k,$$

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

$$(DM_R(\theta))_{ij} = -\operatorname{Cov}_{R,\theta}(X^{2i}, \psi_j(X)).$$

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 fourteenth 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.

[17] Anonymous, *The critical signal fraction for local entropy concavity*, Evidence Press, 2026, doi:10.5281/zenodo.23910629.

[18] A. Eskenazis, P. Nayar and T. Tkocz, Sharp comparison of moments and the log-concave moment problem, *Adv. Math.* **334** (2018), 389–416. `https://arxiv.org/abs/1801.07597`.

[19] M. Fradelizi and O. Guédon, A generalized localization theorem and geometric inequalities for convex bodies, *Adv. Math.* **204** (2006), 509–529. `https://doi.org/10.1016/j.aim.2005.05.020`.

[20] X. Chen, S. He, B. Jiang, C. T. Ryan and T. Zhang, The discrete moment problem with nonconvex shape constraints, arXiv:1708.02079. https://arxiv.org/abs/1708.02079.

[21] Y. Eitan, The centered convex body whose marginals have the heaviest tails, arXiv:2110.14382. https://arxiv.org/abs/2110.14382.