Local and eventual entropy concavity under Gaussian smoothing

Scott Schubr

Abstract

For independent copies of a symmetric variance-one input, we study entropy of their variance-preserving weighted sum after Gaussian smoothing with retained signal fraction q. For every $Q < 1/2$, we prove all-weight strict entropy concavity, simultaneously for $0 < q \le Q$, in an explicit infinite-dimensional Gaussian-relative L^∞ neighborhood. We also prove a common all-weight threshold for symmetric all-moment families with a uniform moment envelope and a finite nonzero Hermite discrepancy. Consequently, among symmetric variance-one log-concave inputs, any counterexample sequence with $q \downarrow 0$ must converge weakly and in every fixed absolute moment to the Gaussian. Each fixed symmetric non-Gaussian all-moment input has eventual all-weight strict concavity, and the whole log-concave class has strictly negative endpoint curvature for $0 < q \le 1/20$. An exact moment-matching exponential-tail construction delineates the remaining near-Gaussian gap. No common all-weight threshold for the full symmetric log-concave class is established.

1 Introduction and results

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

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

Given independent copies U_q, V_q of X_q, set

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

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

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

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

The entropy-concavity question for log-concave inputs originates with Ball, Nayar and Tkocz; we use its formulation and historical attribution as given in Luo [1]. Luo constructs asymmetric counterexamples for every positive q and symmetric, strongly log-concave counterexamples arbitrarily close to the Gaussian for $q > 4/7$. The symmetric low-signal range remains a separate question. Gaussian scale mixtures form a class with entropy concavity [5]. Abu-Ajamieh [2] gives an exact symmetric endpoint identity, several endpoint estimates, and an all-weight result for a particular fourth-Hermite perturbation family. The Hermite-coordinate and Gaussian-channel tools used here have classical antecedents [3, 4].

Our first result is a uniform local theorem for arbitrary bounded even relative perturbations, with infinitely many permitted Hermite modes. 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.

Theorem 1.1 (Uniform Gaussian-relative neighborhood)**.** *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,*

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

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

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

Thus $F_X(q,\cdot)$ is strictly concave unless $g = 0$ almost everywhere. The input need not be smooth or log-concave.

An explicit sufficient radius is as follows:

$$\begin{aligned} c &= \frac{1 + (2Q)^{-1/2}}{2}, & r &= \frac{c-1}{4}, & \rho &= \frac{1+2r}{c}, \\ K &= \frac{\sqrt{\rho}}{(1-\sqrt{\rho})^2}, & A &= \frac{K}{8} + 60 + \frac{1}{8r^2}, & \delta_Q &= \min\left\{\frac{1}{12}, \frac{8}{K}, \frac{1}{16A}\right\}. \end{aligned} \tag{3}$$

These constants are not optimized. Their degeneration as $Q \uparrow 1/2$ is part of the scope of the result.

Corollary 1.2. *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.*

This answers the counterexample-sequence question in Luo [1, Problem 10.2] negatively for $0 < q < 1/2$.

The local hypothesis is substantial: Gaussian-smoothed Laplace inputs have unbounded density ratio to φ for every $q > 0$. Our second result allows these tails, but changes the order of quantifiers.

Theorem 1.3 (Eventual concavity for each input)**.** *Let X be symmetric about zero, with $\mathbb{E}X^2 = 1$ and $\mathbb{E}|X|^m < \infty$ for every integer $m \ge 1$. If X is non-Gaussian, 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. If $n = 2k \ge 4$ is the first order at which the moments of X differ from Gaussian moments and κ_n is its nonzero nth cumulant, then*

$$F_X(q,t) = h(\varphi) - \frac{\kappa_n^2}{2n!} q^n \left[t^k + (1-t)^k\right]^2 + o(q^n) \tag{4}$$

in $C^2([0,1])$ as $q \downarrow 0$.

In particular this holds for every fixed symmetric variance-one log-concave input, and also for discrete symmetric inputs with all moments finite. No numerical threshold uniform over those inputs follows.

A finite-discrepancy refinement, proved in Section 5, gives a common threshold on all symmetric all-moment families with a common moment envelope and a positive discrepancy in finitely many Hermite moments. Its consequence for the log-concave class is that every sequence of counterexamples with signal fractions tending to zero must approach the Gaussian weakly and in every fixed absolute moment. Section 7 constructs exactly moment-matched exponential-tail laws in that remaining boundary region; their entropy curvature is not assigned a sign.

Theorem 1.4 (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*

$$\partial_t^2 F_X(q,0) = \partial_t^2 F_X(q,1) \le -\frac{1}{128}\mathbb{E}\left[\left(1 + (\log p)''(X_q)\right)^2\right]. \tag{5}$$

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.

2 A Gaussian-gain regularity lemma

The following argument applies under the hypotheses of Theorem 1.3.

Lemma 2.1. *Let X, Y be independent copies of a symmetric variance-one random variable with all absolute moments finite. For $H = (h_1, h_2)$ sufficiently near zero put*

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

Then K is smooth near zero and even in each coordinate. The function

$$J(a,b) = K(\sqrt{a}, \sqrt{b}), \qquad a, b \ge 0,$$

is smooth up to both axes on a sufficiently small positive quadrant.

Proof. Work on a closed gain disk with σ bounded away from zero. The density is

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

For each finite multiindex α, Gaussian-kernel differentiation gives

$$\frac{\partial_H^\alpha p_H(z)}{p_H(z)} = \mathbb{E}[P_{\alpha,H}(X, Y, G) \mid Z_H = z], \tag{6}$$

where $P_{\alpha,H}$ is a polynomial with coefficients bounded on the disk. This follows by replacing the standardized kernel residual by G in the joint construction. On compact z intervals differentiation under the expectation follows from finite input moments and bounded Gaussian-kernel derivatives.

Choose R with $\Pr(|X| \le R, |Y| \le R) > 0$. Restricting the kernel expectation to this event yields uniform constants

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

Every formal derivative of $\int p_H \log p_H$ is a finite sum of integrals of a density derivative times $1 + \log p_H$, or of p_H times products of ratios from (6). Conditional Hölder and Jensen inequalities bound these terms in each fixed $L^s(p_H)$ by finitely many moments of X, Y, G, uniformly on the disk. The logarithm is handled by its quadratic bound and moments of Z_H. Higher moments and Hölder's inequality make the integral tails over $|Z_H| > L$ vanish uniformly as $L \to \infty$. On compact spatial intervals all terms are continuous with positive denominators. Differentiation under the entropy integral to every fixed order, and continuity of the derivatives, follow by this compact-interval and uniform-tail decomposition. The same proof applies to $D(p_H \| \varphi)$ because $\log \varphi$ is quadratic.

Symmetry gives evenness in each gain. To handle square roots at the axes, if E is a smooth even function and $a = h^2$, then for $j \geq 1$

$$\frac{d^j}{da^j} E(\sqrt{a}) = \frac{1}{2^{2j-1}(j-1)!} \int_0^1 (1-u^2)^{j-1} E^{(2j)}(uh)\, du. \tag{7}$$

For $j = 1$ this is the fundamental theorem of calculus applied to $E'(h)$; induction follows by differentiation and integration by parts. The formula extends continuously to $h = 0$, with value $j!E^{(2j)}(0)/(2j)!$. Applying it in both coordinates proves all the asserted mixed-derivative extensions. Taylor formulas on the quadrant follow by integrating along segments contained in it. □

3 Uniform local theorem below one-half signal

We prove Theorem 1.1. Fix $Q < 1/2$ and the constants (3), and put $\delta = \|g\|_\infty \leq \delta_Q$. The conditions on g imply that its Hermite coefficients a_n vanish except at even orders $n \geq 4$. The output density from (1) is $\varphi(1 + R)$, where

$$R = L + M, \qquad L = T_{\sqrt{qt}} g + T_{\sqrt{q(1-t)}} g, \tag{8}$$

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 $\|R\|_\infty \leq 2\delta + \delta^2 \leq 3\delta =: \eta \leq 1/4$ on the real weight interval. The output has mass one and variance one, so

$$F_{tt} = -\langle R'', \log(1 + R)\rangle - \int \frac{(R')^2}{1+R}\varphi. \tag{9}$$

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

3.1 Linear energy

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

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

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

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

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

$$\langle L'', L\rangle \geq D^2/8, \qquad \left\|N^{-1}L''\right\|_2 \leq D, \qquad \|L\|_2 \leq D/4. \tag{10}$$

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

3.2 Bilinear term and a smoothing margin

Gaussian convolution in Hermite coordinates [3] gives

$$M_n(q,z) = q^{n/2} \sum_{j+k=n} a_j a_k \binom{n}{j}^{1/2} z^{j/2}(1-z)^{k/2}. \tag{11}$$

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

$$\begin{gathered} P \leq q(1+2r) < \bar{P} < 1, \qquad P/\bar{P} \leq \rho < 1, \\ \bar{P}^2 = c^2 q^2 \leq c^2 Q q < q/2. \end{gathered} \tag{12}$$

Absolute coefficient majorization in (11) reduces the norm to the real bilinear expression $B(P,w; g_{\mathrm{abs}}, g_{\mathrm{abs}})$. For any $h \in L^2(\varphi)$,

$$\|B(P,w;h,h)\|_2 \leq \left\|T_{\sqrt{P}} h\right\|_2^2. \tag{13}$$

Indeed smooth two independent inputs separately and condition their product on their weighted Gaussian sum; L^2 contraction and independence give (13). Polynomial approximation extends the identity and inequality from finite Hermite sums 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 (12) and $S_k(t) \geq 2^{1-k}$,

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

Consequently $\|M(q,z)\|_2 \leq \delta \|L\|_2 /2$ 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 (12) 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 (10) yield

$$\|M\|_2 \leq \delta D/8, \qquad \|NM\|_2 \leq K\delta D/8, \qquad \|M''\|_2 \leq \delta D/(4r^2). \tag{14}$$

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 $0 < q \leq Q$.

3.3 A bounded-function estimate

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

$$\left\|(u_x)^2\right\|_2 \le 3\,\|u\|_\infty\,\|Nu\|_2\,. \tag{15}$$

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

$$\|N\psi(R)\|_2 \le \left[\frac{\eta}{1-\eta} + \frac{3\eta}{(1-\eta)^2}\right]\|NR\|_2 \le 30\delta(1 + K\delta/8)D. \tag{16}$$

3.4 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 (9): 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 (16).

Drop the last, nonpositive term of (9) 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 1.1. Corollary 1.2 follows by taking $Q = q$ and observing that every sufficiently large member of the proposed sequence satisfies its relative-density hypothesis.

4 Eventual all-weight concavity for each input

We now prove Theorem 1.3. The Gaussian moment sequence is determinate: its even moments obey $(2m-1)!! \le (2m)^m$, so the Carleman sum diverges. A non-Gaussian symmetric all-moment input therefore has a first differing even moment $n = 2k \ge 4$. Its cumulants of orders $3, \dots, n-1$ vanish, and $\kappa_n = \mathbb{E}X^n - \mathbb{E}G^n \ne 0$.

Use the gain model from Lemma 2.1. 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

$$p_{[n]}(H,z) = \frac{\kappa_n}{n!}(h_1^n + h_2^n)\varphi(z)\,\mathrm{He}_n(z). \tag{17}$$

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

Write $D(H) = D(p_H\|\varphi)$. By Lemma 2.1, all its finite derivatives 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 (17). Since $\int \mathrm{He}_n^2\,\varphi = n!$,

$$D(H) = \frac{\kappa_n^2}{2n!}(h_1^n + h_2^n)^2 + R_0(H). \tag{18}$$

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 (7) identifies Taylor coefficients at the origin and gives

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

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

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

uniformly over positive-quadrant directions, including the axes. Substituting $a = qt, b = q(1-t)$ proves (4) in $C^2([0,1])$: two weight derivatives introduce $q^2(R_{aa} - 2R_{ab} + R_{bb})$.

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

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

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

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

for all $0 < q \le q_X$ and all weights. This proves Theorem 1.3. Log-concave laws have exponential tails, hence all absolute moments, giving the stated corollary for that class.

5 Uniform restoration away from the Gaussian

The fixed first-nonzero-moment argument alone does not control changing lower cumulants. A finite entropy-jet estimate supplies a stronger uniform statement.

Theorem 5.1 (Finite Hermite discrepancy)**.** *Let $\mathcal{C}$ be a family of symmetric variance-one laws with a uniform absolute-moment envelope: for each integer $m \geq 1$,*

$$\sup_{X \in \mathcal{C}} \mathbb{E}|X|^m \leq B_m < \infty.$$

Put $a_n(X) = \mathbb{E}e_n(X)$. Suppose that for some even $N \geq 4$ and $\varepsilon > 0$,

$$\sum_{\substack{4 \leq n \leq N \\ n \text{ even}}} a_n(X)^2 \geq \varepsilon^2 \qquad (X \in \mathcal{C}). \tag{19}$$

There is $q_{\mathcal{C},N,\varepsilon} > 0$ such that every $X \in \mathcal{C}$ has strict all-weight entropy concavity for every $0 < q \leq q_{\mathcal{C},N,\varepsilon}$, including continuous second derivatives at both endpoints. No input density or log-concavity is required.

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 2.1 is uniform, as are the density upper bound and quadratic logarithm bound. Its conditional-polynomial derivative formulas and the common moment envelope bound every fixed-order entropy derivative uniformly. Formula (7) transfers these bounds to squared-gain derivatives on a common small quadrant. If $P_{X,N}$ denotes the degree-N Taylor polynomial of J_X, then

$$|\partial^\alpha (J_X - P_{X,N})(a,b)| \leq C_N (a+b)^{N+1-|\alpha|}, \qquad |\alpha| \leq 2. \tag{20}$$

Here and below constants may depend on finitely many B_m and on N. The integral Taylor formula applied to the derivatives uses uniform derivatives through degree $N+1$. 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

$$\frac{1}{2} \sum_{\substack{4 \leq n \leq N \\ n \text{ even}}} a_n(X)^2 (a^{n/2} + b^{n/2})^2. \tag{21}$$

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

$$\left(\prod_{\ell=1}^{m} a_{n_\ell}(X) \right) a^r b^s, \quad m \geq 3, \quad n_\ell \geq 4 \text{ even}, \quad r + s = \frac{1}{2} \sum_{\ell=1}^{m} n_\ell \leq N. \tag{22}$$

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 (22) 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, and their remaining power of q is at most q^2, because there is at least one remaining index and each index is at least four. 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 (20)–(21) 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 (19), $E_{X,N}(q) \ge \varepsilon^2 q^N$. 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. □

Corollary 5.2 (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 $p \ge 2$ [6] 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 [8], 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 (19). Theorem 5.1 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. □

As a special case, fixing a first differing moment order and a positive cumulant gap gives a common threshold. Theorem 5.1 also permits lower moments to vary and lower cumulants to vanish. Its constants are not uniform as the finite discrepancy tends to zero; Corollary 5.2 leaves inputs approaching the Gaussian as the remaining boundary.

6 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 3.

Gaussian convolution preserves log-concavity, and the posterior-variance identity for the Gaussian kernel gives

$$0 \le g \le M := \frac{1}{1-q}. \tag{23}$$

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

$$F_{tt}(q, 0) = -\frac{A}{2} - \frac{\kappa_4(U)}{12}B, \qquad B = \mathbb{E}g^2 - \frac{1}{3}\mathbb{E}\rho^4. \tag{24}$$

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 2.1 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 (24). 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 [7]. 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

$$K \le 1 + (1+2C)A \le 116/95 < 5/4. \tag{25}$$

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 $p \ge 2$ [6] 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 (25). Combining the two bounds gives $|\kappa_4(U)B| < (189/32)A$. Equation (24) 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 1.4.

7 Moment-matched exponential tails

The Gaussian-relative and weak/moment neighborhoods in the preceding results are distinct. The following elementary construction, based on a tangent-tail potential and compactly supported moment corrections, makes their difference explicit. It is a limitation of combining the stated hypotheses, not an entropy-concavity counterexample.

Theorem 7.1. *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. □

8 Scope and remaining question

The restoration statements have different quantifiers. Theorem 1.1 supplies a common explicit interval of signal fractions inside a relative-density neighborhood, with that neighborhood depending on its upper signal bound. Theorem 1.3 covers every symmetric all-moment input, but its threshold may depend on the law. Theorem 5.1 supplies uniformity for positive finite Hermite discrepancies, and Corollary 5.2 confines a hypothetical small-signal counterexample sequence to the weak Gaussian boundary. Theorem 1.4 has a common signal interval for the entire symmetric log-concave class, but only controls endpoint curvature. These results retain the stated law-dependent, neighborhood-dependent and endpoint quantifiers.

The obstruction to simply combining them is genuine at the level of hypotheses. Theorem 7.1 gives exact finite Gaussian moment matching together with exponential tails, convergence to the Gaussian, and first unmatched cumulants tending to zero. Such inputs stay outside the bounded relative-density neighborhood after every positive retained signal. This is a limitation of the local topology. Changing inputs with increasingly weak or high-order non-Gaussian moments therefore remain a relevant boundary for the full-class question. This boundary motivates relative-curvature estimates and the analysis of concavity on the moment-matched family.

References

[1] C. Luo, *Near-Gaussian counterexamples to the Ball–Nayar–Tkocz entropy concavity conjecture*, arXiv:2609.11418v2, 2026. `https://arxiv.org/abs/2609.11418v2`.

[2] F. Abu-Ajamieh, *Noise-Induced Restoration of Entropy Concavity Near Gaussian Equilibrium*, arXiv:2610.22386v1, 2026. `https://arxiv.org/abs/2610.22386v1`.

[3] E. Abbe and L. Zheng, *A coordinate system for Gaussian networks*, arXiv:1002.2271, 2010. `https://arxiv.org/abs/1002.2271`.

[4] 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`.

[5] A. Eskenazis and L. Gavalakis, *On the entropy and information of Gaussian mixtures*, Mathematika, 2024, doi:10.1112/mtk.12246.

[6] D. Murawski, *Comparing moments of real log-concave random variables*, arXiv:2211.05210. `https://arxiv.org/abs/2211.05210`.

[7] P. Cattiaux, A. Guillin and P.-A. Zitt, *Poincaré inequalities and hitting times*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, doi:10.1214/11-AIHP447; see equation (4.1). `https://www.numdam.org/item/10.1214/11-AIHP447.pdf`.

[8] 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`.