Self-concordance and allocation for spatial power sums

Mathematical supplement to
Dimension Thresholds for Concavity of Normalized Mixed Moments

Scott Schubr
Luminary Oddities

October 24, 2026
Archive edition 3

Abstract

We study self-concordance of the negative logarithm of one-sided spatial power-sum utilities. For binary generating polynomials whose logarithms are concave on the nonnegative orthant and whose coefficients do not increase when an element is added, one absolute scaling factor C_* makes the original potential a standard self-concordant barrier at every real power $p > 3$, independently of dimension and support rank. Independent coordinate powers $p_i > 3$ are also allowed, with barrier parameter at most $3C_* \sum_i p_i$. The strict threshold is sharp for a uniform theorem with global C^3 regularity. Matroid independent-set polynomials satisfy the structural premise by established log-concavity results. For arbitrary decreasing continuous or lattice inputs, a power linear in dimension is sufficient; sparse lattice support replaces dimension by the number of positive coordinates in each state. Coefficient deletion ratios up to M admit powers of order $(1+\log M)^2$, and a rank-one obstruction shows this asymptotic order is necessary for uniform convexity. For an explicitly listed m-state input, a balanced start has objective gap at most $\log m$, enabling classical damped Newton minimization. For the quartic matroid utility, a balanced allocation box and finite-variance median-of-means estimates give an independence-oracle allocation guarantee with sample dependence linear in ϵ^{-1}, up to logarithmic factors. The optimization and sampling methods are classical; the verified thresholds and uniform resource bounds concern these particular utilities.

1 Introduction and statements

The central fixed-objective result is Theorem 1.1: for structured binary inputs, one absolute scaling retains every real power $p > 3$, and the threshold is sharp in the global C^3 definition. Corollary 7.1 gives the corresponding statement for independently chosen coordinate powers. Scaling the potential preserves its minimizer; changing its smoothing power changes the utility. For the larger class of arbitrary decreasing inputs, the dimension and sparsity bounds instead control how large a common power must be.

The allocation results distinguish two access models. A listed input allows exact moment evaluation and damped Newton steps. For the implicit quartic matroid utility, a balanced allocation box and median-of-means gradient estimates give a sampled allocation algorithm using an independence oracle. Its proof uses curvature and second moments and does not require a self-concordance premise. Expected oracle cost and success probability are stated separately.

The companion manuscript [1] proves log-concavity of spatial product-power smoothing for decreasing inputs, using conditional residual moments and a covariance estimate. This supplement

controls third derivatives of the same raw potential and studies its self-concordance. The companion's sharp leading log-concavity thresholds are distinct from the sufficient power bounds below. An arbitrary convex negative logarithm need not have the self-concordance needed for the usual affine-invariant Newton analysis. Indeed, the elementary quadratic example in Section 6 fails every fixed positive self-concordance scaling.

For a C^3 convex function g on an open convex domain, standard self-concordance means

$$|D^3 g(A)[h,h,h]| \le 2\big(D^2 g(A)[h,h]\big)^{3/2}. \tag{1}$$

A barrier additionally tends to $+\infty$ at every finite boundary point. Its parameter is at most ν if

$$|Dg(A)[h]| \le \sqrt{\nu D^2 g(A)[h,h]}. \tag{2}$$

These definitions, the scaling and affine restriction rules, and the Newton results used below are classical [2, 3].

We consider either

$$F_p(A) = \int_{[0,\infty)^d} \mu(I) \prod_{i=1}^{d} (A_i - I_i)_+^p \, dI \tag{3}$$

or

$$F_p(A) = \sum_{n \in \mathbb{Z}^d_{\ge 0}} \mu(n) \prod_{i=1}^{d} (A_i - n_i)_+^p, \qquad A \in (0,\infty)^d. \tag{4}$$

For (3), assume that μ is nonnegative, coordinatewise nonincreasing, integrable on every finite box, and has positive integral on some box. For (4), assume that μ is finite-valued, nonzero, nonnegative and coordinatewise nonincreasing. The total continuous mass and the global array support may be infinite. The dimension, sparsity and activity bounds use a common real power. Corollary 7.1 permits independent coordinate powers for the structured binary class.

For a binary input, the generating polynomial exposes an additional structure. The following theorem is the fixed-objective result; the dimension and sparsity statements below treat the larger class of arbitrary decreasing inputs.

Theorem 1.1 (Structured inputs above cubic power)**.** *Let $Q(z) = \sum_{S \subseteq [d]} a_S \prod_{i \in S} z_i$ be multiaffine with nonnegative coefficients, $a_\varnothing > 0$, and $a_{S \cup \{i\}} \le a_S$ for $i \notin S$. Suppose $\log Q$ is concave on the whole nonnegative orthant. For every common real $p > 3$, set*

$$F_p(A) = \prod_i A_i^p \, Q\big(((1 - 1/A_i)_+^p)_i\big), \qquad g = -\log F_p.$$

Then $C_* g$, for the absolute constant $C_* = 2560000$, is a standard self-concordant barrier with parameter at most $3C_* pd$. The same scale works uniformly for all $p > 3$, dimensions, support ranks and coefficients. For $p \ge 4$ the parameter bound $2C_* pd$ is retained. The strict threshold is sharp for a uniform theorem under the global C^3 definition: for every $0 < p \le 3$, $Q(z) = 1 + z$ already produces a potential that is not C^3 at its activation point. If Q is constant, the unscaled separable barrier already works for $p \ge 1$.

Theorem 1.2 (Dimension bound)**.** *For either input class, if $p \ge 16d$, then $g = -\log F_p$ is a standard self-concordant barrier on the positive orthant. Its parameter is at most*

$$\nu_d = \frac{d(p + 6/5)^2}{p - 16d/5} \le 2pd.$$

Theorem 1.3 (Sparse support bound)**.** *Suppose the input in (4) is supported on states with at most r positive coordinates, where $1 \le r \le d$. If $p \ge 32r$, then $g = -\log F_p$ is a standard self-concordant barrier with parameter at most*

$$\nu_r = \frac{d(p + 6/5)^2}{p - 16r/5} \le 2pd.$$

There is no bound on the sizes of the nonzero integer coordinates or on the number of states. For $r = 0$, the function is the separable barrier $-p \sum_i \log A_i - \log \mu(0)$, with parameter pd for $p \ge 1$.

For weighted independent sets of a rank-r matroid, with coordinate activities in $[0, 1]$, the input satisfies Theorem 1.3. The sparse theorem itself requires no matroid hypothesis or negative association. The sparse theorem's parameter still depends on d: only the required *power* is controlled by r. Theorem 1.1 gives an alternative for log-concave binary generating polynomials: scaling by an absolute constant retains every fixed power $p > 3$, including the quartic utility, and admits independent coordinate powers. A classical orthant logarithmic barrier has parameter d. The results here control the particular raw smoothing potential, whose optimization objective one may wish to retain.

Classical context

The entropic barrier of Bubeck and Eldan [4] is a Fenchel dual of a log-Laplace transform; its self-concordance proof uses third moments of log-concave measures. The universal barrier studied by Lee and Yue [5] is a log polar-volume potential, with sharp moment inequalities for marginals of convex bodies. Chewi [11] proves the exact dimension parameter for the entropic barrier. Neither identification is assumed for (3)–(4): their inputs can be non-log-concave and atomic. Generalized self-concordance [6] provides broader third-order hypotheses and Newton analysis, not an implication from a Hessian sign alone. Infimal-convolution smoothing, as in [7], is a different operation from the integral or sum here. Hyperbolic-polynomial barriers [10] also have different structural hypotheses; we do not assume real stability or hyperbolicity of a power sum. Our contribution is the residual-moment verification for this specific smoothing class, and its sparse-support refinement.

2 Residual tails and exact derivatives

Every active box in (3) is bounded, and every active sum in (4) is finite. Monotonicity and nonzero input ensure $F_p(A) > 0$ throughout the orthant. In the continuous case a positive lower box lies below every positive A; in the discrete case $\mu(0) > 0$. For $p > 3$ the kernel x_+^p is C^3, and its first three derivatives are bounded on bounded intervals. Differentiation of the original fixed-input formulas on compact positive A sets is therefore justified by local integrability or local finiteness, including at activation walls.

At a fixed A, normalize the active integral or sum to a probability law, put $z_i = A_i - I_i$ or $A_i - n_i$, and define

$$Z_i = p(A_i/z_i - 1) \ge 0. \tag{5}$$

Zero residuals carry zero tilted mass. The residual representation is used only at the fixed point; its A-dependent weight is not differentiated.

Lemma 2.1 (Residual moments)**.** *For either input class and $s \ge 0$,*

$$\Pr\{Z_i > s\} \le (1 + s/p)^{-p}. \tag{6}$$

For $p \geq 8$ this implies

$$\mathbb{E}Z_i \leq M_1 = 6/5, \qquad \mathbb{E}Z_i^2 \leq M_2 = 16/5, \qquad \mathbb{E}Z_i^3 \leq M_3 = 15. \tag{7}$$

The coordinates may be dependent.

Proof. Condition on all other coordinates. In the continuous case the residual density is proportional to $z^p w(z)$ on $(0, A_i)$, with w nondecreasing. For $0 < u < A_i$, the integrals below and above u are bounded respectively by

$$\frac{w(u)u^{p+1}}{p+1}, \qquad \frac{w(u)(A_i^{p+1} - u^{p+1})}{p+1}.$$

Their ratio gives $\Pr\{z < u\} \leq (u/A_i)^{p+1}$; truncation or monotone approximation covers zero and unbounded endpoint weights.

In the discrete case the positive residuals are an arithmetic list $t, t+1, \ldots, A_i$ and their weights are nondecreasing. Match the k entries below u in order to the largest k entries. The weight does not decrease and the residual ratio is at least A_i/u. Thus lower-block mass is at most $(u/A_i)^p$ times upper-block mass, which is at most the total mass. The argument only compares sums and remains valid if the blocks overlap. Setting $u = A_i/(1+s/p)$ and integrating the conditional bounds proves (6).

Tail integration gives, for $k = 1, 2, 3$ and $p > k$,

$$\mathbb{E}Z_i^k \leq \frac{k! p^k}{(p-1)\cdots(p-k)}.$$

The right sides decrease with p, and at $p = 8$ are respectively $8/7$, $64/21$ and $512/35$. These are bounded by (7). □

Fix a direction h and write $a_i = h_i/A_i$,

$$X = \sum_i a_i Z_i, \qquad Y = \sum_i a_i^2 (2Z_i + Z_i^2/p). \tag{8}$$

For the logarithm ℓ of one active product kernel,

$$\ell' = p \sum_i a_i + X, \qquad \ell'' = -p \|a\|_2^2 - Y,$$
$$\ell''' = 2 \sum_i a_i^3 (p + 3Z_i + 3Z_i^2/p + Z_i^3/p^2).$$

Partition differentiation yields the exact identities

$$g'' = p \|a\|_2^2 + \mathbb{E}Y - \operatorname{Var} X, \tag{9}$$
$$g''' = -\mathbb{E}\ell''' + 3 \operatorname{Cov}(X, Y) - \mathbb{E}(X - \mathbb{E}X)^3. \tag{10}$$

These follow by differentiating the original partition three times and collecting the first three cumulants. The moment bounds justify the identities at walls as well. In particular, the covariance and third cumulant in (10) must not be discarded.

3 Proof of the dimension bound

By Cauchy–Schwarz,

$$\operatorname{Var} X \le M_2 \|a\|_1^2 \le \frac{16d}{5} \|a\|_2^2, \qquad g'' \ge (p - 16d/5) \|a\|_2^2. \tag{11}$$

Hölder's inequality gives $\mathbb{E}(Z_i Z_j^2) \le M_3$ and $\mathbb{E}Z_i \mathbb{E}Z_j^2 \le M_1 M_2$, while $|\operatorname{Cov}(Z_i, Z_j)| \le M_2$. Therefore

$$\begin{aligned} 3|\operatorname{Cov}(X,Y)| &\le 3\left(2M_2 + \frac{M_3 + M_1 M_2}{p}\right) \|a\|_1 \|a\|_2^2 \\ &< 27\sqrt{d}\, \|a\|_2^3. \end{aligned} \tag{12}$$

Indeed at $p = 8$ the bracket with its factor 3 is $5253/200 < 27$. Since $|z - m|^3 \le z^3 + m^3$ for nonnegative z, m,

$$\mathbb{E}|Z_i - \mathbb{E}Z_i|^3 \le M_3 + M_1^3 = 2091/125 < 17.$$

Minkowski's inequality, with arbitrary dependence, gives

$$|\mathbb{E}(X - \mathbb{E}X)^3| \le 17 \|a\|_1^3 \le 17 d^{3/2} \|a\|_2^3. \tag{13}$$

Also

$$|\mathbb{E}\ell'''| \le \left(2p + 6M_1 + \frac{6M_2}{p} + \frac{2M_3}{p^2}\right) \|a\|_2^3 < (2p + 11) \|a\|_2^3. \tag{14}$$

Combining (10)–(14) and $d \ge 1$ yields

$$|g'''| \le (2p + 55 d^{3/2}) \|a\|_2^3.$$

For $p \ge 16d$, division by the curvature lower bound gives

$$\begin{aligned} \frac{|g'''|}{(g'')^{3/2}} &\le \frac{2/\sqrt{p} + 55(d/p)^{3/2}}{(1 - 16d/(5p))^{3/2}} \\ &\le \frac{1/2 + 55/64}{(4/5)^{3/2}} = \frac{435\sqrt{5}}{512} < 2. \end{aligned}$$

The last comparison is the integer inequality $946125 < 1048576$. This proves (1) and strict convexity.

The first derivative obeys

$$|g'[h]| \le (p + M_1) \|a\|_1 \le (p + 6/5)\sqrt{d}\, \|a\|_2,$$

which with (11) gives ν_d. Since $p - 16d/5 \ge 4p/5$ and $(p + 6/5)/p \le 43/40$, $\nu_d \le 1849pd/1280 < 2pd$. At any finite orthant boundary point one coordinate A_i tends to zero and the others stay bounded. Local integrability bounds $F_p(A)$ by a constant times A_i^p in (3); a fixed finite active box gives the same bound for (4). Thus $F_p \to 0$ and $g \to +\infty$, completing Theorem 1.2.

4 Sparse support and the rank bound

Lemma 4.1 (Sparse L^3 operator)**.** *Suppose $Z_i \geq 0$, $\mathbb{E}Z_i^3 \leq M$, and at most r of the Z_i are nonzero on each outcome. Then, for every real vector a,*

$$\left\| \sum_i a_i Z_i \right\|_{L^3} \leq M^{1/3} \sqrt{r} \, \|a\|_2 .$$

Proof. Take b with $\|b\|_{L^{3/2}} \leq 1$ and put $B_i = \mathbb{E}[|b|^{3/2} \mathbf{1}_{Z_i > 0}]$. Then $0 \leq B_i \leq 1$ and $\sum_i B_i \leq r$. Hölder gives $|\mathbb{E} Z_i b| \leq M^{1/3} B_i^{2/3}$, so

$$\left| \mathbb{E} b \sum_i a_i Z_i \right| \leq M^{1/3} \|a\|_2 \left(\sum_i B_i^{4/3} \right)^{1/2} \leq M^{1/3} \sqrt{r} \, \|a\|_2 .$$

The L^3 dual supremum proves the assertion. □

For the sparse lattice law, $Z_i > 0$ precisely when $n_i > 0$ on its tilted support. Pointwise Cauchy–Schwarz gives $X^2 \leq r \sum_i a_i^2 Z_i^2$, hence

$$g'' \geq (p - 16r/5) \, \|a\|_2^2 . \tag{15}$$

Lemma 4.1 gives $\|X\|_{L^3} \leq 15^{1/3} \sqrt{r} \, \|a\|_2$, so centering costs at most a factor two and

$$|\mathbb{E}(X - \mathbb{E}X)^3| \leq 120 r^{3/2} \, \|a\|_2^3 .$$

Minkowski and the monotonicity of L^q norms give

$$\|Y\|_{L^{3/2}} \leq (2 \cdot 15^{1/3} + 15^{2/3}/p) \, \|a\|_2^2 .$$

Consequently

$$\begin{aligned} 3|\operatorname{Cov}(X, Y)| &\leq 6 \, \|X\|_{L^3} \, \|Y\|_{L^{3/2}} \\ &\leq (12 \cdot 15^{2/3} + 90/p) \sqrt{r} \, \|a\|_2^3 < 85 \sqrt{r} \, \|a\|_2^3 . \end{aligned}$$

Here $p \geq 8$ and $15^{2/3} < 49/8$; the latter follows by cubing, $115200 < 117649$. With (14), for $r \geq 1$,

$$|g'''| \leq (2p + 216 r^{3/2}) \, \|a\|_2^3 .$$

For $p \geq 32r$ its ratio to (15) to the power $3/2$ is at most

$$\frac{2/\sqrt{32} + 216/32^{3/2}}{(9/10)^{3/2}} = \frac{175\sqrt{5}}{216} < 2,$$

since $153125 < 186624$. The gradient estimate is unchanged, giving ν_r. Since the denominator is at least $9p/10$ and $p \geq 32$, $\nu_r \leq pd(83/80)^2/(9/10) < 2pd$. Boundary divergence and C^3 regularity were already established. The case $r = 0$ is immediate from $\mu(0) \prod_i A_i^p$. This proves Theorem 1.3.

The sparsity hypothesis counts positive coordinates, not their magnitudes. It therefore includes unbounded occupancies along a sparse collection of axes. A nonzero ordinary Lebesgue density cannot be supported on a union of lower-dimensional coordinate subspaces.

5 A linear-order obstruction for arbitrary decreasing inputs

The companion [1] already proves dimension-linear lower obstructions, including unweighted hereditary binary supports, and a sharp leading dimension coefficient for its broader spatial classes. The following mixture is a convenient explicit smooth-cell witness for the order comparison here.

Proposition 5.1. *For every real $p \geq 1$ and every integer $d \geq 64(p+1)$, there is a finite decreasing binary input for which $-\log F_p$ is not convex at an interior point. Thus a uniform convexity, and hence self-concordance, theorem for arbitrary decreasing sparse binary inputs requires powers of order their maximal active-coordinate cardinality.*

Proof. Set $\mu(0) = 1$, $\mu(n) = q$ at every other binary state, and $\mu = 0$ outside $\{0,1\}^d$, with $0 < q < 1$. This is coordinatewise decreasing and has exact partition

$$F_p(A) = (1-q)\prod_i A_i^p + q\prod_i (A_i^p + (A_i - 1)_+^p).$$

Take $A_i = p+1$ and put

$$b = \left(\frac{p}{p+1}\right)^p, \qquad t = \frac{b}{1+b}, \qquad q = \frac{1}{1+(1+b)^d}.$$

The two components have equal tilted mass. The first component has zero occupancy; the other has independent Bernoulli(t) occupancies. The residual Z_i is now the occupancy indicator. In direction $h_i = A_i$, if K is total occupancy,

$$\mathbb{E}K = dt/2, \quad \operatorname{Var} K = dt(1-t)/2 + d^2t^2/4, \quad Y = (2+1/p)K.$$

Equation (9) gives

$$g'' = d\left[p + \frac{t}{2}(1+1/p+t)\right] - \frac{d^2t^2}{4}.$$

Since $\log(1+1/p) < 1/p$, we have $b > e^{-1} > 1/3$, and therefore $1/4 < t < 1/2$. Also $\frac{t}{2}(1+1/p+t) < 1$ for $p \geq 1$. Thus

$$g'' < d(p+1) - d^2/64 \leq 0.$$

The curvature is strictly negative even at the displayed endpoint. For $p \leq 3$ this is still a legitimate smooth-cell computation, since all the selected residuals are strictly positive. □

For an explicit rational witness, take $p = 4$, $A_i = 2$ and $q = 16^d/(16^d+17^d)$. The two components again have equal tilted mass, and the product component has Bernoulli$(1/17)$ occupancies with $Z_i = 4n_i$. In direction $h_i = 2$,

$$\mathbb{E}Y = 6d/17, \qquad \operatorname{Var} X = (128d + 4d^2)/289, \qquad g'' = (1130d - 4d^2)/289.$$

At $d = 283$ the last quantity is exactly $-566/289$. The verifier checks this by differentiating the literal two-product expression, without enumerating 2^{283} states. Padding with coordinates forced to zero occupancy preserves the example in its original directions, so the obstruction is in active-coordinate cardinality r, not necessarily ambient dimension. This weight is generally not a product-activity matroid law; no such obstruction for that narrower class is inferred.

6 The scalar fourth-power repair

The quadratic function $F(x) = x^2 + (x-1)_+^2$ gives a useful failed endpoint. On $x > 1$ put $P = 2x^2 - 2x + 1$. Then

$$(-\log F)'' = \frac{8x(x-1)}{P^2}, \qquad (-\log F)''' = \frac{8(2x-1)(-2x^2+2x+1)}{P^3}.$$

As $x \downarrow 1$, curvature tends to zero while the third derivative tends to 8. Hence no fixed positive scaling is self-concordant, even inside that smooth chamber.

Proposition 6.1. *For every $0 \le q \le 1$, $-\log[x^4 + q(x-1)_+^4]$ is standard self-concordant on $(0,\infty)$. For every real $p \ge 4$, the function $-100\log[x^p + q(x-1)_+^p]$ is standard self-concordant, uniformly in q.*

Proof. For the quartic active chamber let $P = x^4 + q(x-1)^4$ and define

$$A = P'^2 - PP'', \qquad C = -P'''P^2 + 3PP'P'' - 2P'^3.$$

The second and third derivatives of $-\log P$ are A/P^2 and C/P^3. The substitution $x = 1+v$, $q = w/(1+w)$, $v, w \ge 0$, clears denominators $(1+w)^2$ and $(1+w)^6$ in A and $4A^3 - C^2$. Their numerator polynomials have respectively 21 and 133 nonzero strictly positive integer coefficients, with positive constant terms. These exact polynomial identities and every coefficient are supplied by the standalone verifier. At $q = 1$ the substituted polynomials in v likewise have strictly positive coefficients. This proves the quartic inequality for all active points and both activity endpoints. On $x \le 1$ the potential is $-4\log x$, and C^3 regularity glues the chambers.

For real $p \ge 4$, put $y = (x-1)/x$, $a = (1-y)/y$ on $x > 1$ and let $Z = pa$ with probability $w = qy^p/(1+qy^p)$ and zero otherwise. Then

$$\frac{x^2 g''}{p} = 1 + 2wa - (p-1)wa^2 + pw^2a^2.$$

Since $w \le y^p$,

$$(p-1)wa^2 \le \frac{4(p-1)}{p^2}(1-2/p)^{p-2} \le \frac{p-1}{p^2} \le \frac{3}{16}.$$

Indeed $(1-2/p)^{p-2}$ decreases for $p \ge 4$ and is $1/4$ at four; the derivative of its logarithm is $\log(1-2/p) + 2/p \le 0$. Thus $g'' \ge 13p/(16x^2)$. For $k = 1, 2, 3$,

$$\mathbb{E}Z^k \le p^k \max_{0 \le y \le 1} y^{p-k}(1-y)^k \le k^k.$$

The scalar form of (10) bounds $x^3|g'''|$ by

$$2p\mathbb{E}(1+Z/p)^3 + 3|\operatorname{Cov}(Z, 2Z + Z^2/p)| + |\mathbb{E}(Z - \mathbb{E}Z)^3| \le 6p + 52 + 81/p \le 25p.$$

Here $\mathbb{E}Z \le 1$, $\mathbb{E}Z^2 \le 4$, $\mathbb{E}Z^3 \le 27$; the three terms are at most $6p$, $24 + 81/p$ and 28. Consequently $|g'''| \le 20(g'')^{3/2}$ for $p \ge 4$, as is checked by squaring at the worst bound $p = 4$. Multiplication by 100 divides the ratio by ten. The inactive chamber and C^3 gluing are unchanged. □

7 Structured inputs and the sharp power threshold

We prove Theorem 1.1. The coefficient deletion and ordinary log-concavity hypotheses permit a uniform third-moment bound, so the original power-sum utility can be retained at every power allowed by the global C^3 definition.

Proof. Write $L = \log Q$ and $\phi_i(A_i) = (1 - 1/A_i)_+^p$. Decomposing $Q = R_i + z_i Q_i$, coefficient deletion gives $Q_i \le R_i$, so $0 \le L_i \le 1/(1+z_i) \le 1$ and $\nabla^2 L \preceq 0$. On an active coordinate put $y_i = 1 - 1/A_i$. The compensated chain rule is

$$g''[h,h] \ge \sum_i \left(\frac{p}{A_i^2} - \phi_i'' L_i \right) h_i^2.$$

The positive part of $A_i^2 \phi_i''/p$ is at most

$$H(p) = \frac{4(p-1)}{p^2}(1 - 2/p)^{p-2} \le H(3) = \frac{8}{27}.$$

Indeed $A_i^2 \phi_i''/p = y_i^{p-2}[(p-1)(1-y_i)^2 - 2y_i(1-y_i)]$, and maximizing its first positive term gives the displayed bound. Moreover $(\log H)' = 1/(p-1) + \log(1-2/p) < 0$ for $p \ge 3$. Inactive coordinates contribute p/A_i^2. With $a_i = h_i/A_i$ this yields

$$g''[h,h] \ge \frac{19p}{27} \|a\|_2^2. \tag{16}$$

For $p \ge 4$, the sharper value $H(4) = 3/16$ also gives $g'' \ge 13p\|a\|_2^2/16$, which will be used for quartic allocation.

At a fixed A, use the binary law proportional to $a_S \phi^S$. Let $I_i = \mathbf{1}_{i \in S}$ and, on active coordinates, $\alpha_i = p/(A_i - 1)$ and $Z_i = \alpha_i I_i$; set $Z_i = 0$ elsewhere. Deletion gives $\Pr(I_i = 1) \le \phi_i = y_i^p$ and, for $j = 1, 2, 3$,

$$\mathbb{E} Z_i^j \le p^j \max_{0 \le y \le 1} y^{p-j}(1-y)^j \le j^j.$$

Given the other indicators equal to J, the conditional occupation probability is at most $\phi_i/(1+\phi_i) \le \phi_i$. If $a_J = 0$, deletion makes the extended coefficient zero too; the conditioning event is then null. Iteration gives

$$\Pr(T \subseteq S) \le \prod_{i \in T} \phi_i. \tag{17}$$

No independence assumption is involved.

Use X, Y and the exact cumulant identity (10). The preceding moments imply

$$|\mathbb{E}\ell'''| \le 9p\|a\|_2^3, \qquad \|Y\|_{3/2} \le (6 + 9/p)\|a\|_2^2 \le 9\|a\|_2^2. \tag{18}$$

For the first bound, $2p + 6 + 24/p + 54/p^2 \le 9p$ for $p \ge 3$.

Normalize $\|a\|_2 = 1$ and put $\gamma_i = a_i \alpha_i$. For the low-score set $|\gamma_i| \le 4$, exponential change of these fields and the tangent inequality for the concave function $\log Q$ give

$$\log \mathbb{E} e^{t(X_{\mathrm{lo}} - \mathbb{E}X_{\mathrm{lo}})} \le \sum_{\mathrm{lo}} m_i (e^{t\gamma_i} - 1 - t\gamma_i), \qquad m_i = \Pr(I_i = 1).$$

Whole-orthant concavity permits either sign of t, even if a changed field exceeds one. Since $\sum_i m_i \gamma_i^2 \le 4$, the inequality $e^u - 1 - u \le 3u^2/2$ for $|u| \le 1$ bounds each log moment at $t = \pm 1/4$ by $3/8$. Both exponential moments are below two. Consequently

$$\|X_{\mathrm{lo}} - \mathbb{E}X_{\mathrm{lo}}\|_3^3 < 1536 < (35/3)^3.$$

On the high-score set $|\gamma_i| > 4$, monotonicity of $p\log(1+\alpha/p)$ gives the cubic bound

$$\phi_i = (1+\alpha_i/p)^{-p} \le (1+\alpha_i/3)^{-3} \le 27\alpha_i^{-3}.$$

Thus for $j = 1, 2, 3$,

$$A_j := \sum_{\text{hi}} \phi_i |\gamma_i|^j \le 27\, 4^{j-3} \sum_i |a_i|^3 \le 27\, 4^{j-3}.$$

Expanding $V = \sum_{\text{hi}} |\gamma_i| I_i$ and using (17),

$$\mathbb{E}V^3 \le A_3 + 3A_2A_1 + A_1^3 \le \frac{270243}{4096} < (25/6)^3.$$

Centering costs at most two in L^3, so Minkowski and homogeneity give

$$\|X - \mathbb{E}X\|_3 < 20\|a\|_2 \qquad (a \neq 0). \tag{19}$$

The split only estimates already evaluated derivatives and is not differentiated. Only moments through order three have been used.

Hölder and (18) bound the covariance contribution by $3 \cdot 20 \cdot 9\|a\|_2^3$. Hence (10) gives

$$|g'''[h,h,h]| \le (9p + 8540)\|a\|_2^3, \qquad 8540 = 20^3 + 3 \cdot 20 \cdot 9.$$

It follows that

$$\frac{|g'''|}{(g'')^{3/2}} \le \frac{9p + 8540}{(19p/27)^{3/2}} \le \frac{8567}{(19/9)^{3/2}} < 3200.$$

The last squared comparison is $8567^2 \cdot 729 < 3200^2 \cdot 19^3$. Scaling by $C_* = 1600^2$ supplies the standard constant two. Also $|g'[h]| \le (p+1)\sqrt{d}\|a\|_2$, giving parameter before scaling

$$\frac{27d(p+1)^2}{19p} \le \frac{48}{19}pd < 3pd.$$

For $p \ge 4$ the sharper curvature bound above retains the upper parameter $2pd$ before scaling. Positivity, boundary divergence and C^3 gluing follow from the finite binary partition when $p > 3$.

For sharpness take $Q(z) = 1 + z$ in dimension one. Near $x = 1$,

$$-\log[x^p + (x-1)_+^p] + p\log x = -(x-1)_+^p + o(|x-1|^p).$$

The left side vanishes for $x \le 1$. If it were C^3, its first three derivatives at one would vanish, and Taylor's theorem would make it $o(|x-1|^3)$. This contradicts the expansion for $0 < p \le 3$. Thus the threshold is necessary for a uniform theorem with the stated global C^3 definition, including after positive scaling. □

Corollary 7.1 (Independent coordinate powers for structured inputs)**.** *Under the coefficient deletion and whole-orthant log-concavity hypotheses of Theorem 1.1, let $p_i > 3$ be independent real powers and set*

$$F(A) = \prod_i A_i^{p_i} Q\left(((1 - 1/A_i)_+^{p_i})_i\right), \qquad g = -\log F.$$

Then $C_* g$ is a standard self-concordant barrier with parameter at most $3C_* \sum_i p_i$, where the same absolute constant $C_* = 2560000$ works for all dimensions, coefficients and power vectors.

Proof. At a fixed point write $a_i = h_i/A_i$ and $\phi_i = (1 - 1/A_i)_+^{p_i}$. On active coordinates put $\alpha_i = p_i/(A_i - 1)$ and $Z_i = \alpha_i I_i$ under the law proportional to $a_S \phi^S$; put $Z_i = 0$ on inactive coordinates. For the logarithm ℓ of one active product kernel,

$$\ell' = \sum_i p_i a_i + X, \qquad X = \sum_i a_i Z_i,$$
$$\ell'' = -\sum_i p_i a_i^2 - Y, \qquad Y = \sum_i a_i^2 (2Z_i + Z_i^2/p_i),$$
$$\ell''' = 2\sum_i a_i^3 (p_i + 3Z_i + 3Z_i^2/p_i + Z_i^3/p_i^2).$$

Differentiating the original finite partition gives the same cumulant identity

$$g''' = -\mathbb{E}\ell''' + 3\,\mathrm{Cov}(X, Y) - \mathbb{E}(X - \mathbb{E}X)^3.$$

All single-coordinate and joint-inclusion estimates of the above-cubic proof hold coordinatewise for $p_i > 3$. In particular $\mathbb{E}Z_i^j \le j^j$ for $j = 1, 2, 3$ and $\phi_i \le 27\alpha_i^{-3}$ on active coordinates. The compensated chain rule and the curvature loss bound $8/27$ give, with $S = \sum_i p_i a_i^2$,

$$g''[h, h] \ge \frac{19}{27} S.$$

The low/high score argument of Theorem 1.1 only uses those uniform moments, cubic tails, joint inclusion and concavity of $\log Q$. Therefore it still gives

$$\|X - \mathbb{E}X\|_3 < 20\|a\|_2, \qquad \|Y\|_{3/2} \le 9\|a\|_2^2.$$

Also, coordinatewise,

$$2p_i + 6 + 24/p_i + 54/p_i^2 \le 9p_i, \qquad |\mathbb{E}\ell'''| \le 9\sum_i p_i |a_i|^3.$$

Since $p_i > 3$, $\|a\|_2 \le \sqrt{S/3}$ and $\max_i |a_i| \le \sqrt{S/3}$. Hence the kernel term is at most $9S^{3/2}/\sqrt{3}$. The centered cubic is at most $8000S^{3/2}/(3\sqrt{3})$, and Hölder's inequality bounds the covariance term by $3 \cdot 20\|a\|_2 \cdot 9\|a\|_2^2 \le 540S^{3/2}/(3\sqrt{3})$. Thus

$$|g'''[h, h, h]| \le \frac{8567}{3\sqrt{3}} S^{3/2}, \qquad \frac{|g'''|}{(g'')^{3/2}} \le \frac{8567}{(19/9)^{3/2}} < 3200.$$

The same scaling $C_* = 1600^2$ supplies the standard constant two.

Finally $\mathbb{E}Z_i \le 1$ and weighted Cauchy–Schwarz give

$$|g'[h]| \le \sum_i (p_i + 1)|a_i| \le \left(\sum_i \frac{(p_i + 1)^2}{p_i}\right)^{1/2} \sqrt{S}.$$

The parameter before scaling is consequently at most

$$\frac{27}{19}\sum_i \frac{(p_i + 1)^2}{p_i} \le \frac{48}{19}\sum_i p_i < 3\sum_i p_i.$$

The finite partition is positive; $p_i > 3$ gives C^3 gluing at every activation wall, and a vanishing coordinate forces F to zero at every finite orthant boundary point. This proves the barrier claim. □

For matroid independent sets with activities $\lambda_i \in [0,1]$, $a_S = \mathbf{1}_{S \text{ independent}} \prod_{i\in S} \lambda_i$ has the deletion property. Ordinary Q log-concavity is classical: Anari–Liu–Oveis Gharan–Vinzant [8, Theorem 4.1] prove complete log-concavity of

$$\sum_{S \text{ independent}} y^{d-|S|} \prod_{i\in S} z_i.$$

The homogenizing degree is the *ground-set size* d, not rank. Restricting to $y = 1$ and inserting nonnegative activities gives exactly the premise used here.

In Corollary 9.1, scaling by C_* preserves the fixed-p utility maximizer. Its starting gap is at most $C_* \log m$, so the classical Newton count is $O(1 + C_* \log m + \log(1 + \log(1/\epsilon)))$ at every $p > 3$ for this structured binary class. Since C_* is absolute, its hidden constant is independent of rank. The guarantee is still for an explicit state list, not an implicit matroid counting oracle.

8 Large activities and squared-logarithmic powers

The deletion constant one cannot be removed while retaining a fixed power. Increasing the common power gives a different sufficient theorem for the same smoothing construction.

Theorem 8.1. *Let $Q(z) = \sum_S a_S z^S$ be multiaffine with nonnegative coefficients, $a_\varnothing > 0$, and $\log Q$ concave on the whole nonnegative orthant. Suppose $a_{S+i} \le M a_S$ for $i \notin S$, where $1 \le M < \infty$. Put $\ell = \log M$ and $H = 1 + \ell$. For every common real power $p \ge 72H^2$, the function*

$$g(A) = -\log\left[\prod_i A_i^p Q(((1 - 1/A_i)_+^p)_i)\right]$$

has Cg a standard self-concordant barrier, with $C = 360000$ and parameter at most $3Cpd$. Weighted matroid independent-set polynomials with activities in $[0, M]$ satisfy these hypotheses by classical ordinary log-concavity [8] and heredity.

Proof. At fixed A, put $\phi_i = (1 - 1/A_i)_+^p$ and tilt S with weights $a_S\phi^S$. On active coordinates let $\alpha_i = p/(A_i - 1)$ and $Z_i = \alpha_i \mathbf{1}_{i\in S}$; elsewhere set $Z_i = 0$. Conditioning on all other indicators gives

$$\Pr(i \in S \mid S \setminus \{i\} = J) \le \frac{M\phi_i}{1 + M\phi_i} \le \beta_i := \min(1, M\phi_i).$$

If $a_J = 0$, the numerator also vanishes; null conditioning events cause no exception. Iteration gives $\Pr(T \subset S) \le \prod_{i\in T} \beta_i$. This is an external-field bound. Both Z_i and the artificial variable $\alpha_i B_i$ with $B_i \sim \text{Bernoulli}(\beta_i)$ have tail bounded by $\min(1, M(1 + t/p)^{-p})$, because $\phi_i = (1 + \alpha_i/p)^{-p}$. This bound is exactly the tail of $b + cY$, where $\Pr(Y > t) = (1 + t/p)^{-p}$, $c = M^{1/p}$ and $b = p(c - 1)$. Here $c < 3$, $b \le 3\ell$, since $\ell/p \le 1$. For $1 \le k \le 4$,

$$\mathbb{E}Y^k = \frac{k!p^k}{(p-1)\cdots(p-k)}.$$

These moments decrease with p; at $p = 8$ they are $8/7, 64/21, 512/35, 4096/35$, whose roots are bounded respectively by $4/3, 2, 8/3, 10/3$. Minkowski therefore gives, for both variables,

$$\|Z_i\|_1 \le 4H, \quad \|Z_i\|_2 \le 6H, \quad \|Z_i\|_3 \le 8H, \quad \|Z_i\|_4 \le 10H. \tag{20}$$

The artificial Bernoulli bound is needed for the joint estimates below.

For $u_i = h_i/A_i$, set $X = \sum_i u_i Z_i$ and $Y_0 = \sum_i u_i^2(2Z_i + Z_i^2/p)$. The original partition identities of Section 2 apply. Write $B = \log Q$ and $\pi_i = \phi_i B_i = \Pr(i \in S)$. On an active coordinate,

$$A_i^2 \phi_i'' B_i = \pi_i[(p-1)\alpha_i^2/p - 2\alpha_i] \le \mathbb{E} Z_i^2.$$

The compensated chain rule and $\nabla^2 B \preceq 0$ imply

$$g''[h,h] \ge (p - 36H^2) \|u\|^2 \ge (p/2) \|u\|^2 . \tag{21}$$

This uses the actual $\phi_i B_i$, not the generally false large-activity bound $B_i \le 1$. Inactive coordinates have separable curvature. Also

$$\|Y_0\|_{3/2} \le (16H + 64H^2/p) \|u\|^2 \le 17H \|u\|^2 ,$$
$$|\mathbb{E}\ell'''_{\text{kernel}}| \le [2p + 24H + 216H^2/p + 1024H^3/p^2] \|u\|^3 \le 3p \|u\|^3 .$$

For the last comparison the extra terms divided by p are at most $1/3 + 1/24 + 1024/72^3 < 1$.

Normalize $\|u\| = 1$ and put $c_i = u_i \alpha_i$. For the low set $|c_i| \le 12H$, concavity of $\log Q$ applied to fields $\phi_i e^{tc_i}$ yields, for either sign of t,

$$\log \mathbb{E} e^{t(X_{\text{lo}} - \mathbb{E}X_{\text{lo}})} \le \sum_{\text{lo}} \pi_i (e^{tc_i} - 1 - tc_i).$$

All fields remain in the whole-orthant concavity domain. Since $\sum_i \pi_i c_i^2 \le 36H^2$, at $t = \pm(12H)^{-1}$ this is at most $3/8$, using $e^v - 1 - v \le 3v^2/2$ for $|v| \le 1$. Both exponential moments are less than two. Hence $|z|^3 \le 6(12H)^3 e^{|z|/(12H)}$ gives

$$\|X_{\text{lo}} - \mathbb{E}X_{\text{lo}}\|_3^3 < 24(12H)^3 < 35^3 H^3.$$

For the high set $|c_i| > 12H$, the artificial fourth moment gives $\beta_i \alpha_i^4 \le (10H)^4$. Thus

$$A_k := \sum_{\text{hi}} \beta_i |c_i|^k \le (10H)^4 (12H)^{k-4} \quad (k = 1,2,3), \qquad A_1 \le 6H,\ A_2 \le 70H^2,\ A_3 \le 834H^3,$$

since $\sum_i |u_i|^4 \le 1$. Expanding the cube of $V = \sum_{\text{hi}} |c_i| \mathbf{1}_{i \in S}$ and using joint inclusion gives $\mathbb{E}V^3 \le A_3 + 3A_2A_1 + A_1^3 \le 2310H^3 < 14^3 H^3$. Centering costs at most two in L^3, so homogeneity yields $\|X - \mathbb{E}X\|_3 < 63H \|u\|$. The exact third-cumulant identity now implies

$$|g'''[h,h,h]| \le [3p + 253260H^3] \|u\|^3 , \qquad 253260 = 63^3 + 3 \cdot 63 \cdot 17,$$

where $H \ge 1$ absorbs the covariance term's H^2. Consequently

$$\frac{|g'''|}{(g'')^{3/2}} \le \frac{3p + 253260H^3}{(p/2)^{3/2}} \le \frac{253476}{216} < 1200.$$

Each summand decreases with p; the last comparison uses $p = 72H^2$ and $H \ge 1$. Scaling by 600^2 gives the standard constant two. Finally $|g'[h]| \le (p + 4H)\sqrt{d} \|u\|$, so the gradient parameter before scaling is at most $2d(p + 4H)^2/p \le 2pd(19/18)^2 < 3pd$. The finite binary sum is positive, is C^3 across walls, and tends to zero at every finite orthant boundary. This proves the barrier claim. □

Proposition 8.2 (A rank-one obstruction)**.** *For each fixed $0 < \gamma < 1/4$, powers $p = \gamma(\log M)^2$ do not give a uniform convexity theorem as $M \to \infty$, even for $Q(z) = 1 + Mz$. At $p = 4, M = 625, A = 5/4$, one has $[-\log(A^4 + 625(A-1)^4)]'' = -192/25$.*

Proof. Set $\ell = \log M$, $y = e^{-\ell/p}$, $A = (1-y)^{-1}$ and $\alpha = p(e^{\ell/p} - 1)$. The two product terms have equal tilted mass, so occupation is Bernoulli 1/2. Direct differentiation gives, for $p > 2$,

$$A^2 g''(A) = p + \alpha - \frac{p-2}{4p}\alpha^2.$$

For $p = \gamma\ell^2$, $\alpha/\ell \to 1$; division by ℓ^2 gives limit $\gamma - 1/4 < 0$. In the finite example $y = 1/5$, $\alpha = 16$, $A^2 g'' = -12$, and therefore $g'' = -192/25$. □

This establishes squared-logarithmic order necessity in the stated asymptotic regime, not a sharp finite-activity threshold. Changing the power changes the objective. In particular it does not extend the quartic sampled allocation theorem to activities above one.

9 An explicit budget-allocation consequence

Corollary 9.1. *Let a finite decreasing lattice input be explicitly listed with m nonzero states and maximal active-coordinate cardinality r. Assume $p \geq 32r$ for $r \geq 1$, or $p \geq 1$ for $r = 0$. For $c_i > 0$ and $B > 0$, maximize $F_p(A)$ subject to $A_i \geq 0$ and $\sum_i c_i A_i = B$. There is a unique maximizer $A^* > 0$. Starting at*

$$A_i^0 = \frac{B}{dc_i},$$

classical damped Newton minimization of $-\log F_p$ on the affine simplex attains $F_p(A) \geq e^{-\epsilon} F_p(A^*)$ in

$$O(1 + \log m + \log(1 + \log(1/\epsilon)))$$

iterations for $0 < \epsilon \leq 1/2$, with an absolute hidden constant. Each iteration uses $O(md^2 + d^3)$ arithmetic operations and $O(md)$ positive-real-power oracle calls.

Proof. Since $\mu(n) \leq \mu(0)$ and every residual is bounded by its endpoint, weighted AM–GM gives

$$F_p(A) \leq m\mu(0) \prod_i A_i^p \leq m\mu(0) \frac{(B/d)^{pd}}{\prod_i c_i^p} \leq m F_p(A^0).$$

Thus the initial objective gap is at most $\log m$. The closed simplex is compact, the partition is positive at A^0 and zero whenever a coordinate vanishes. Its maximizer is therefore interior. Strict convexity of the negative logarithm on the affine slice proves uniqueness. For $d = 1$ or $m = 1$ the displayed point is already optimal.

In affine coordinates on the tangent space, the restriction f is a nondegenerate standard self-concordant barrier. Its Newton decrement is $\delta^2 = \nabla f^{\mathsf{T}} (\nabla^2 f)^{-1} \nabla f$. The classical damped step [3, Lecture 2, IX–X]

$$x^+ = x - \frac{(\nabla^2 f)^{-1} \nabla f}{1 + \delta}$$

remains feasible and decreases the objective by at least $\delta - \log(1+\delta)$. While $\delta \geq 1/4$ the decrease is at least $1/4 - \log(5/4) > 0$, so there are $O(1 + \log m)$ such steps. Once $\delta \leq 1/4$, the same classical theorem gives

$$\delta_+ \leq 2\delta^2, \qquad f(x) - f^* \leq \frac{\delta^2}{2(1-\delta)}.$$

The remaining count is $O(1 + \log(1 + \log(1/\epsilon)))$. Additive error in $-\log F_p$ is precisely the claimed multiplicative utility accuracy. Summing each listed state's gradient and Hessian costs $O(md^2)$, and a dense tangent or bordered Newton solve costs $O(d^3)$. Product rules or prefix/suffix products avoid division by a zero residual at a wall. □

This is an application of classical Newton theory to the stated explicit state-list model. The guarantee concerns this specific utility and retains its objective, rather than replacing it by the standard orthant logarithmic barrier.

10 Sampled allocation for an implicit matroid

The preceding Newton consequence uses an explicit state list. For the specific quartic matroid utility, sampling instead gives an implicit independence-oracle guarantee. Projected gradient descent and matroid sampling are classical; the point here is the uniform balanced box for this utility.

Theorem 10.1. *Let $\mathcal{I}$ be the independent sets of a matroid on $[d]$, accessible by an independence oracle of cost t_I. Let $0 \le \lambda_i \le 1$, $c_i > 0$ and $B > 0$, and put*

$$F(A) = \sum_{S \in \mathcal{I}} \prod_{i \in S} \lambda_i \prod_{i=1}^{d} (A_i - \mathbf{1}_{i \in S})_+^4.$$

For $0 < \epsilon, \delta \le 1/2$, there is a randomized algorithm producing $A > 0$, $c \cdot A = B$, with $-\log F(A) \le \min_{c \cdot A' = B}[-\log F(A')] + \epsilon$ with probability at least $1-\delta$. Its expected number of arithmetic operations and independence-oracle calls is polynomial in $d, \epsilon^{-1}, \log(1/\delta)$, independently of the smallest positive activity, c and B, in the exact real-arithmetic sampling model. The number of sampler calls is

$$O\left(\frac{d}{\epsilon}[1 + \log(d/\epsilon)] \log \frac{d[1 + \log(d/\epsilon)]}{\delta}\right).$$

Every history produces a feasible output. No partition-value, counting, or Hessian oracle, bit-complexity bound, or self-concordance premise is needed for this algorithm.

Proof. Write $g = -\log F$. At a positive A, the tilted independent-set law has fields $w_i = \lambda_i(1 - 1/A_i)_+^4$. Set $Z_i = 4/(A_i - 1)$ on occupation of coordinate i, and zero otherwise (in particular when $A_i \le 1$). The deletion estimates and the compensated log-concavity calculation of Section 7 give, with $m_i = \mathbb{E} Z_i$ and $D_A = \operatorname{diag}(A_i)$,

$$0 \le m_i \le 1, \quad \mathbb{E} Z_i^2 \le 4, \qquad \partial_i g = -(4 + m_i)/A_i,$$
$$\frac{13}{4} I \preceq D_A \nabla^2 g D_A = 4I + \operatorname{diag} \mathbb{E}(2Z_i + Z_i^2/4) - \operatorname{Cov}(Z) \preceq 7I.$$

The upper matrix bound uses $\operatorname{Cov}(Z) \succeq 0$.

The utility is continuous on the compact budget simplex, positive in its interior and zero on its coordinate boundary. Strict convexity of g therefore gives a unique interior minimizer A^*. The Lagrange multiplier equations for the original budget constraint give

$$\frac{4 + m_i(A^*)}{A_i^*} = \tau c_i, \qquad \tau B = 4d + \sum_i m_i(A^*).$$

Set $A_i^0 = B/(dc_i)$, $x_i = A_i/A_i^0$ and $h(x) = g(A^0 x)$. Then

$$x^* \in K := \{x \in [4/5, 5/4]^d : \sum_i x_i = d\}, \qquad \mu I \preceq \nabla^2 h(x) \preceq LI \quad (x \in K), \quad \mu = \frac{52}{25}, \quad L = \frac{175}{16}.$$

Moreover $\nabla h(x^*)$ is parallel to $(1, \dots, 1)$, even if x^* is on a face of this artificial box. Thus for every $x \in K$,

$$h(x) - h(x^*) \le (L/2) \|x - x^*\|^2. \tag{22}$$

The Euclidean projection $P_K(u)$ is exactly $[P_K(u)]_i = \min(5/4, \max(4/5, u_i - \theta))$, with θ making the sum d. Sort the $2d$ breakpoints $u_i - 5/4, u_i - 4/5$ and solve the piecewise affine sum on its appropriate interval; a flat interval may use either endpoint. This costs $O(d \log(2d))$ arithmetic operations. The projection variational inequalities give $\|P_K(u) - P_K(v)\|^2 \le (P_K(u) - P_K(v)) \cdot (u - v)$, hence nonexpansiveness. Integrating the Hessian on a segment in K shows that $x \mapsto x - \nabla h(x)/L$ has Lipschitz constant $q = 1 - \mu/L = 3543/4375$. Also $P_K(x^* - \nabla h(x^*)/L) = x^*$. Consequently the iteration

$$x_{j+1} = P_K(x_j - \widehat{v}_j/L), \quad x_0 = \mathbf{1}, \qquad \|\widehat{v}_j - \nabla h(x_j)\| \le \eta$$

satisfies $\|x_T - x^*\| \le q^T \sqrt{d}/4 + \eta/\mu$.

Choose the largest power of two $\eta \le 1$ satisfying $\eta^2 \le \mu^2 \epsilon/(2L)$, and the largest power of two $\zeta \le 1$ satisfying $\zeta^2 \le 16\eta^2/(25d)$. Let $T \ge 1$ be the smallest integer with $q^{2T} d \le 8\epsilon/L$. As above, these choices use rational comparisons and successive multiplication, and accurate gradients at all T steps imply $h(x_T) - h(x^*) \le \epsilon$.

We estimate the score means by the classical median-of-means method [12, Section 2.1]. Set

$$a = \left\lceil 64\zeta^{-2} \right\rceil, \qquad b = \min\{j \ge 1 : j \text{ is odd and } 2^j \ge 4dT/\delta\}, \qquad N = ab, \qquad \xi = \frac{\delta}{4TN}.$$

At each iterate make N fresh, mutually independent calls to the weighted independent-set sampler, each within total variation ξ of the current tilted law. Divide the draws into b groups of size a. For each coordinate let $\widehat{m}_i$ be the median of its b group means of the unmodified score Z_i, and use

$$\widehat{v}_i = -\frac{4 + \widehat{m}_i}{x_i}.$$

The score is zero on inactive or zero-field coordinates. No clipping or fourth-moment estimate is required.

Here is the conditional error bound at any fixed current iterate, including an iterate reached after previous inaccurate estimates. For a independent exact draws, the previously proved second-moment bound gives

$$\operatorname{Var}\left(\frac{1}{a}\sum_{j=1}^{a} Z_i^{(j)}\right) \le \frac{4}{a}, \qquad \Pr\left\{\left|\frac{1}{a}\sum_{j=1}^{a} Z_i^{(j)} - m_i\right| > \zeta\right\} \le \frac{4}{a\zeta^2} \le \frac{1}{16}.$$

For a fixed coordinate the groups are independent. If its median has error greater than ζ, at least $(b+1)/2$ groups are bad. A union bound over subsets of bad groups therefore gives

$$\Pr\{|\widehat{m}_i - m_i| > \zeta\} \le 2^b (1/16)^{(b+1)/2} \le 2^{-b}.$$

No independence between different coordinates is needed. A further union bound makes the probability of any coordinate error at most $d2^{-b} \le \delta/(4T)$ for exact draws.

The joint law of the N independent approximate draws differs from that of N exact independent draws by total variation at most $N\xi = \delta/(4T)$, by the product-measure telescoping inequality (or a coordinatewise maximal coupling). Applying this bound to the whole median-accuracy event shows that, conditionally on every previous history, the failure probability for the actual batch is at most $\delta/(2T)$. This argument compares event probabilities, not expectations of an unbounded score, so it requires no uniform score bound.

On a successful batch,

$$\|\widehat{v} - \nabla h(x)\| \le \frac{5}{4}\sqrt{d}\,\zeta \le \eta.$$

The deterministic contraction and gap bounds proved above then give accuracy after T steps whenever every batch succeeds. The conditional union bound bounds any batch failure by $\delta/2$, and hence in particular by δ. At every history, including histories with inaccurate batches, projection keeps the iterate in K and the returned budget allocation strictly positive and feasible.

For cost accounting, $\zeta^{-2} = O(d/\epsilon)$ and $T = O(1 + \log(d/\epsilon))$. The total number of sampler calls is

$$TN = O\left(\frac{d}{\epsilon}\,[1 + \log(d/\epsilon)] \log \frac{d[1 + \log(d/\epsilon)]}{\delta}\right).$$

All budgets are fixed before the iterations. Since each positive field is at most one, the sampler of Chen–Guo–Zhang–Zou [9, Theorem 1] has conditional expected cost

$$O(d \log(2d/\xi)(\log(2d) + t_I))$$

per draw, uniformly over every history. At each call delete zero-field elements; their scores are zero, while their separable gradient contribution remains $-4/A_i$. An empty restriction has the deterministic empty-set sample. The fixed iteration count is retained even if a particular restriction is empty. Positive remaining fields may be arbitrarily small. Summing conditional expectations gives that bound times TN for total expected sampling cost. Computing all scores and group means costs $O(dTN)$ further arithmetic operations; sorting the group means to take their medians costs $O(Tdb \log(2b))$, and projections cost $O(Td \log(2d))$. These bounds are polynomial in the parameters stated in the theorem, with sample count linear rather than 4/3-power in ϵ^{-1}, up to logarithmic factors. The real-arithmetic sampling model and absence of a bit-complexity claim are unchanged. □

11 Reproducibility and remaining scope

The accompanying Python 3 verifier uses only the standard library and explicit exceptions, including under optimized execution. It constructs the scalar quartic polynomials by integer coefficient arithmetic, performs the positive-variable substitution, and checks every sign and both activity endpoints. It also differentiates the literal two-product curvature witness with exact rational arithmetic and checks the arithmetic comparisons used in the constant bounds. The resulting JSON records the full coefficient payload. It also checks the fixed-power binary scaling constants. The verifier checks the above-cubic scaling and coordinate-power constants, as well as the exact balanced-box gradient constants, the large-activity constants and its literal scalar curvature witness. These computations support the finite polynomial identities; the arbitrary-dimensional inequalities are proved analytically above.

The constants 16 and 32, the scalar multiplier 100, and the barrier parameters are not sharp. The lower construction establishes the linear order in active support for the unrestricted decreasing class, not a sharp coefficient. The fixed-power structured result has an absolute sufficient scale. The large-activity coefficient 72 and scale 360000 are sufficient; the rank-one obstruction establishes the squared-logarithmic asymptotic order. The structured threshold $p > 3$ is sharp for global C^3 regularity, while its scaling constant is sufficient. The explicit-state Newton result does not furnish an implicit partition oracle; the separate matroid sampling result instead uses an independence oracle and retains the quartic objective, using only finite score variance and median-of-means estimation without a counting or Hessian oracle.

AI assistance. OpenAI ChatGPT/Codex assisted with literature discovery, proof exploration and drafting. The named author is responsible for the mathematical content and attribution.

References

[1] S. Schubr, *Dimension Thresholds for Concavity of Normalized Mixed Moments*, companion manuscript, October 2026, spatial smoothing and binary-structure sections. The present article is a mathematical supplement to that work.

[2] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, 1994. doi:10.1137/1.9781611970791.

[3] A. Nemirovski, *Interior Point Polynomial Time Methods in Convex Programming*, lecture notes, 1996, Lecture 2, Section 2.2, IX–X. Author notes.

[4] S. Bubeck and R. Eldan, *The entropic barrier: a simple and optimal universal self-concordant barrier*, arXiv:1412.1587v3, 2015. arXiv:1412.1587.

[5] Y. T. Lee and M.-C. Yue, *Universal Barrier Is n-Self-Concordant*, Math. Oper. Res. **46** (2021), 1129–1148. arXiv:1809.03011.

[6] T. Sun and Q. Tran-Dinh, *Generalized Self-Concordant Functions: A Recipe for Newton-Type Methods*, arXiv:1703.04599v1, 2017. arXiv:1703.04599.

[7] A. D. Adeoye and A. Bemporad, *Self-concordant smoothing in proximal quasi-Newton algorithms for large-scale convex composite optimization*, arXiv:2309.01781v3, 2025. arXiv:2309.01781.

[8] N. Anari, K. Liu, S. Oveis Gharan and C. Vinzant, *Log-Concave Polynomials III: Mason's Ultra-Log-Concavity Conjecture for Independent Sets of Matroids*, arXiv:1811.01600v1, 2018, Theorem 4.1. arXiv:1811.01600.

[9] X. Chen, H. Guo, X. Zhang and Z. Zou, *Near-Linear Time Samplers for Matroid Independent Sets with Applications*, APPROX/RANDOM 2024, LIPIcs **317**, 32:1–32:12, Theorem 1. doi:10.4230/LIPIcs.APPROX/RANDOM.2024.32.

[10] O. Güler, *Hyperbolic Polynomials and Interior Point Methods for Convex Programming*, Math. Oper. Res. **22** (1997), 350–377. doi:10.1287/moor.22.2.350.

[11] S. Chewi, *The Entropic Barrier Is n-Self-Concordant*, Geometric Aspects of Functional Analysis, Lecture Notes in Mathematics **2327** (2023), 209–222. doi:10.1007/978-3-031-26300-2_6; arXiv:2112.10947v1, 2021.

[12] G. Lugosi and S. Mendelson, *Mean Estimation and Regression Under Heavy-Tailed Distributions: A Survey*, Found. Comput. Math. **19** (2019), 1145–1190, Section 2.1. arXiv:1906.04280.