Heat-capacity annealing for cold-start logconcave sampling

August 2, 2026

Abstract

We study cold-start sampling from a centered isotropic logconcave law $\pi(\mathrm{d}x) \propto e^{-V(x)}\,\mathrm{d}x$ whose finite convex potential V is given by an evaluation oracle. Under the standard ground-set normalization of Kook and Vempala, we construct, for output order q and a suitable internal order $Q = \max\{q, \widetilde{O}(1)\}$, a sample Y satisfying $\mathcal{R}_q(\mathrm{law}(Y) \| \pi) \le \varepsilon$ using

$$\widetilde{O}\Big(d^{5/2}\sqrt{Q} + Qd^2\Big)$$

expected evaluation queries. For fixed or polylogarithmic order this is $\widetilde{O}(d^{2.5})$, improving the $\widetilde{O}(qd^{2.75})$ finite-order cold-start bound of Kook and Vempala on the centered isotropic subclass.

The algorithm is organized by the radial heat capacity of Gaussian annealing,

$$\pi_\beta(\mathrm{d}x) \propto e^{-V(x)-\beta\|x\|^2/2}\,\mathrm{d}x, H(\beta) = \mathrm{Var}_{\pi_\beta}\left(\|X\|^2\right), \qquad H_* = \sup_{\beta \ge 0} H(\beta).$$

For a convex body K with $B(x_0, r) \subseteq K$ and $\pi = \mathrm{Unif}(K)$, the same argument gives the heat-parameterized membership-query bound

$$\widetilde{O}\left(d + \frac{d^2\sqrt{QH_*}}{r^2} + \frac{Qd^2\,\|\mathrm{Cov}(\pi)\|_{\mathrm{op}}}{r^2}\right).$$

Here the convex-body parameter is $H_* = \sup_{\beta\ge 0} \mathrm{Var}_{\pi_\beta} \|X - x_0\|^2$. The polynomial dependence on Q is displayed; the suppressed factors are logarithmic in the geometric parameters and the target accuracy.

We prove $H_* = O(\mathcal{C}_d^3 d)$ for every centered isotropic logconcave law, where $\mathcal{C}_d$ is the worst isotropic logconcave Poincaré constant. Klartag's established $\mathcal{C}_d = O(\log(ed))$ bound already gives $H_* = \widetilde{O}(d)$ and the claimed $d^{2.5}$ exponent. The key evaluation-oracle bridge is that Kook and Vempala's fixed exponential-lift truncation retains at least one half of the mass at every Gaussian precision; conditioning therefore increases the radial heat capacity by at most a factor of two. A recent quadratic-form inequality of Letwin optionally sharpens H_* to $O(d)$, but removes only a hidden polylogarithmic factor and is not needed for either sampling exponent.

Note. This paper was fully AI generated by GPT-5 in the Codex harness, from a ChatGPT Canvas handoff of unspecified model provenance supplied by Yang Liu, and was prompted by Yang Liu.

1 Introduction

Sampling from a logconcave distribution is one of the basic algorithmic tasks at the interface of convex geometry, probability, and optimization. Its membership-oracle form already contains a canonical problem. Given a convex body $K \subset \mathbb{R}^d$, a membership oracle for K, and an accuracy parameter ε, one seeks a random point whose law is close to the uniform distribution π on K. The difficulty is not merely convergence from a favorable initial law. A useful algorithm must also construct such a law from the geometric input—a *cold start*—without paying an extra factor polynomial in the dimension.

The modern oracle-complexity theory grew out of volume computation. Dyer, Frieze, and Kannan gave the first randomized polynomial-time volume algorithm and an underlying near-uniform sampler [8]; Applegate and Kannan initiated the corresponding theory for general logconcave functions [1]. The conductance method of Lovász and Simonovits [23, 24], the localization framework of Kannan, Lovász, and Simonovits [9, 10], and the later geometry of logconcave functions developed by Lovász and Vempala [28] made random-walk sampling a general-purpose primitive. Hit-and-Run, introduced by Smith [30], was shown to mix rapidly by Lovász [22]; the analysis of Lovász and Vempala permits an arbitrary interior starting point and separates this guarantee from warm-start analyses [26]. These algorithms differ in their input normalization, initial-law requirement, output distance, and cost per transition, so their transition bounds and oracle bounds are not interchangeable.

Annealing supplies the usual bridge from a cold input to a warm law. Lovász and Vempala used simulated annealing for convex bodies and logconcave functions [25, 27]. Cousins and Vempala developed Gaussian cooling for volume and Gaussian-volume computation [6, 7]. In a different but related literature, thermodynamic-speed schedules based on heat capacity go back at least to Salamon et al. [29]; adaptive variance and partition-function schedules also have a long theoretical computer science history, notably Štefankovič et al. [31]. Our claim is not that heat-capacity-guided scheduling is itself new. The issue here is whether the particular radial heat capacity of Gaussian cooling can be controlled strongly enough, and coupled to a finite-Rényi membership-oracle sampler, to improve cold-start query complexity.

The proximal-sampling viewpoint of Lee, Shen, and Tian introduced the restricted Gaussian oracle as a structured primitive [20]. Chen, Chewi, Salim, and Wibisono sharpened the analysis of the resulting proximal sampler [5]. Kook, Vempala, and Zhang implemented this diffusion for convex bodies through the In-and-Out chain and obtained strong Rényi guarantees [17, 18]. From a cold start, Kook and Zhang obtained $\widetilde{O}(d^3)$ membership queries for an $\mathcal{R}_\infty$ guarantee [16]. Kook and Vempala then reduced the total-variation warm-start generation cost to $\widetilde{O}(d^2R^{3/2}\Lambda^{1/4})$ [12], where

$$R^2 = \mathbb{E}_\pi \|X - x_0\|^2 , \qquad \Lambda = \|\mathrm{Cov}(\pi)\|_{\mathrm{op}} ,$$

and subsequently obtained the balanced finite-order bound $\widetilde{O}(qd^2R^{3/2}\Lambda^{1/4})$ [15]. In centered isotropic position, $R = \sqrt{d}$ and $\Lambda = 1$, so the latter bound is $\widetilde{O}(qd^{2.75})$, or $\widetilde{O}(d^{2.75})$ for fixed or polylogarithmic order. Their sampler from a warm start already costs $\widetilde{O}(qd^2\Lambda)$; the $qd^{2.75}$ term is the cost of creating the warm law, not of mixing from it. A later unified analysis removes the $\Lambda \vee 1$ loss for arbitrary evaluation-oracle logconcave densities, giving a $\widetilde{O}(qd^2\Lambda)$ finite-order sampling-from-a-warm-start cost under its ground-set normalization [14].

Kook and Vempala also located a genuine obstruction to one tempting acceleration. A spherical Gaussian tilt of an isotropic logconcave law need not preserve the operator norm of the covariance: their counterexample has a tilted covariance eigenvalue $\Omega(d^{1/3})$ [15, Proposition 6.3]. Thus an analysis that charges every phase to the worst direction cannot in general keep the covariance scale constant. This counterexample does not, however, lower-bound the scalar fluctuation

$$\mathrm{Var}_{\pi_\beta} \|X - x_0\|^2 .$$

The present paper exploits exactly this distinction.

1.1 Results

For $\beta \geq 0$, set

$$\pi_\beta(\,\mathrm{d}x) = \frac{1}{Z(\beta)} \exp\!\left(-\frac{\beta}{2}\left\|x - x_0\right\|^2\right) \pi(\,\mathrm{d}x), \qquad H(\beta) = \mathrm{Var}_{\pi_\beta} \left\|X - x_0\right\|^2. \tag{1}$$

Our algorithmic statement first separates this scalar geometry from the sampler.

Theorem 1.1 (Annealing from a heat-capacity bound, informal)**.** *Let $B(x_0, r) \subseteq K$, let π be uniform on K, and suppose $H(\beta) \leq \overline{H}$ for every $\beta \geq 0$. Suppose numerical values x_0, r and upper bounds $\overline{H}$ and $\overline{\Lambda} \geq \|\mathrm{Cov}(\pi)\|_{\mathrm{op}}$ are supplied. Let $q \geq 1$ be the desired output order and let $Q \geq \max\{q, 2\}$ also meet the polylogarithmic order required by the finite-Rényi restart caps. There is a cold-start membership-oracle algorithm whose output Y satisfies $\mathcal{R}_q(\mathrm{law}(Y)\|\pi) \leq \varepsilon$ and whose expected query complexity is*

$$\widetilde{O}\left(d + \frac{d^2\sqrt{Q\overline{H}}}{r^2} + \frac{Qd^2\overline{\Lambda}}{r^2}\right). \tag{2}$$

No additional polynomial dependence on Q is hidden in the annealing term. The $\widetilde{O}$ notation suppresses logarithms in the displayed parameters and in $1/\varepsilon$.

The path is parameterized by precision rather than variance. Its log-partition function $A(\beta) = \log Z(\beta)$ satisfies $A''(\beta) = H(\beta)/4$. Consequently, if $\beta > \beta'$, then

$$\mathcal{R}_Q(\pi_\beta \| \pi_{\beta'}) \leq \frac{Q}{8}(\beta - \beta')^2 \sup_{\theta \in [\beta', \beta' + Q(\beta - \beta')]} H(\theta).$$

An additive mesh of size $\Theta((Q\overline{H})^{-1/2})$ therefore keeps consecutive targets finitely warm. A restricted-Gaussian phase at precision β costs $\widetilde{O}(d^2/(\beta r^2))$ membership queries. Summing these costs over the additive grid produces a harmonic series and the first nontrivial term in (2). The last term is the existing balanced warm-start sampler.

The geometric statement proved here controls the entire path, including the movement of its mean.

Theorem 1.2 (Radial heat capacity)**.** *Let π be any centered isotropic logconcave probability measure on $\mathbb{R}^d$, and define π_β by (1) with $x_0 = 0$. Let*

$$\mathcal{C}_d = \sup\{\mathsf{C}_{\mathrm{P}}(\nu) : \nu \textit{ is centered, isotropic, and logconcave on } \mathbb{R}^d\}.$$

Then

$$\sup_{\beta \geq 0} H(\beta) \leq C\mathcal{C}_d^3 d \tag{3}$$

for a universal constant C. In particular, the established bound $\mathcal{C}_d = O(\log(ed))$ gives $\sup_\beta H(\beta) = O(d\log^3(ed))$.

Using the recent quadratic-form Poincaré inequality of Letwin in place of the general Poincaré inequality sharpens the conclusion to

$$\sup_{\beta \geq 0} H(\beta) = O(d). \tag{4}$$

The distinction between the two conclusions is deliberate. The first part of Theorem 1.2, combined with Klartag's $O(\log(ed))$ Poincaré bound [11], already proves the same $\widetilde{O}(d^{2.5})$ query exponent. The much newer theorem of Letwin [21] removes a geometric polylogarithmic loss but is not needed for the headline dimension exponent. Nor does Letwin's result prove the full KLS conjecture: it gives a dimension-free inequality for quadratic forms and an $O(\log^{1/4}(ed))$ KLS isoperimetric bound, corresponding to an $O(\sqrt{\log(ed)})$ worst Poincaré bound under the conventions used here. We avoid the common notation collision between that paper's third-moment parameter and the worst Poincaré constant.

More quantitatively, in the normalization $r = \lambda = 1$ and with $Q = \widetilde{O}(1)$, the $\widetilde{O}(d^{2.5})$ conclusion needs only $H_* = \widetilde{O}(d)$. Before suppressing logarithmic factors, the annealing term is $d^2\sqrt{QH_*}$; hence the exact condition for it to be $o(d^{2.75})$ is $QH_* = o(d^{3/2})$. In particular, any estimate $H_* = O(d^{3/2-\delta})$ for a fixed $\delta > 0$ gives a genuine polynomial improvement when Q is polylogarithmic. Through $H_* = O(\mathcal{C}_d^3 d)$, the corresponding exact condition is $Q\mathcal{C}_d^3 = o(d^{1/2})$; for example, $\mathcal{C}_d = O(d^{1/6-\delta})$ for any fixed $\delta > 0$ is sufficient. Klartag's logarithmic bound is therefore more than sufficient, and the Letwin-based linear heat bound is an optional polylogarithmic sharpening rather than a dependency of the result.

Combining the dependency-minimal form (3) with Theorem 1.1 gives the main algorithmic corollary.

Corollary 1.3 (Centered scalar-covariance cold start)**.** *Let $K \subset \mathbb{R}^d$ be a convex body given by a membership oracle, suppose $B(x_0, r) \subseteq K$, and assume $x_0 = \mathbb{E}_\pi X$ and $\mathrm{Cov}(\pi) = \lambda I$, where x_0, r, λ are supplied to the algorithm. For every $q \geq 1$ and $\varepsilon \in (0, 1/100)$, choose $Q \geq \max\{q, 2\}$ satisfying the polylogarithmic cap condition of Theorem 3.2. There is an algorithm with*

$$\mathcal{R}_q(\mathrm{law}(Y)\|\pi) \leq \varepsilon$$

using

$$\widetilde{O}\Big(\frac{\lambda}{r^2}(d^{5/2}\sqrt{Q} + Qd^2)\Big) \tag{5}$$

expected membership queries. For fixed-order Rényi sampling, or for total-variation error ε using an internal $Q = \widetilde{O}(1)$, the bound is $\widetilde{O}(\lambda d^{2.5}/r^2)$. In particular it is $\widetilde{O}(d^{2.5})$ when $r = \lambda = 1$.

The same heat estimate accelerates the expensive final phase of the evaluation-oracle construction. Write

$$L_{\pi,g} = \{x \in \mathbb{R}^d : V(x) - \min V \leq 10d\}$$

for the ground set used in that construction.

Theorem 1.4 (Centered isotropic evaluation-oracle cold start)**.** *Let $V : \mathbb{R}^d \to \mathbb{R}$ be a convex function given by an evaluation oracle. Suppose $\pi(\,\mathrm{d}x) \propto e^{-V(x)}\,\mathrm{d}x$ is centered and isotropic and $B(0,1) \subseteq L_{\pi,g}$. For every $q \geq 1$ and $\varepsilon \in (0, 1/100)$, there is a cold-start algorithm returning Y with*

$$\mathcal{R}_q(\mathrm{law}(Y)\|\pi) \leq \varepsilon$$

after

$$\widetilde{O}\Big(d^{5/2}\sqrt{Q} + Qd^2\Big) \tag{6}$$

expected evaluation queries, where $Q = \max\{q, \widetilde{O}(1)\}$ meets the finite-order cap conditions. In particular, the complexity is $\widetilde{O}(d^{2.5})$ for fixed or polylogarithmic output order.

Theorem 1.4 uses the same fixed truncation as Kook and Vempala; it does not replace their exponential lift or their sampler. The additional ingredient in our argument is a uniform comparison between the truncated and untruncated Gaussian paths, described below and proved in Section 6.

The centering and scalar-covariance hypotheses are substantive. A spherical quadratic tilt is not affine invariant, so whitening does not turn an arbitrary anisotropic instance into Theorem 1.2. We do not claim the scale-aware bound $H_* = O(R^2\Lambda)$ for general covariance.

1.2 Proof overview

Write m_β and Σ_β for the mean and covariance of π_β, and introduce

$$G_\beta = \operatorname{tr}(\Sigma_\beta^2) + m_\beta^\top \Sigma_\beta m_\beta, \qquad \Phi_\beta = -\log\det\Sigma_\beta. \tag{7}$$

Energy monotonicity gives $\operatorname{tr}\Sigma_\beta + \|m_\beta\|^2 \le d$, while Brascamp–Lieb gives $\Sigma_\beta \preceq \beta^{-1}I$. A deterministic eigenvalue calculation then yields

$$G_\beta \le d + \frac{3\Phi_\beta}{\beta}. \tag{8}$$

The exact covariance derivative gives

$$\Phi'_\beta = \frac{1}{2}\operatorname{Cov}_{\pi_\beta}\left((X-m_\beta)^\top\Sigma_\beta^{-1}(X-m_\beta), \|X\|^2\right).$$

After whitening π_β, a general Poincaré inequality bounds the two variances in this covariance by $4\mathcal{C}_d d$ and $4\mathcal{C}_d G_\beta$. Hence $\left|\Phi'_\beta\right| \le 2\mathcal{C}_d\sqrt{dG_\beta}$. A first-crossing argument against a linear barrier gives $\Phi_\beta = O(\mathcal{C}_d^2 d\beta)$, then $G_\beta = O(\mathcal{C}_d^2 d)$ and $H(\beta) = O(\mathcal{C}_d^3 d)$.

Letwin's quadratic-form theorem can be inserted at exactly two points. It gives $H(\beta) \le 16G_\beta$ and bounds the whitened Mahalanobis variance by $8d$. The same first-crossing argument then gives $\Phi_\beta = O(d\beta)$ and $H(\beta) = O(d)$. This route is cleaner, but the preceding Poincaré route is retained to show that the sampling exponent is robust to substantially weaker geometry.

1.3 General evaluation-oracle densities

For $\pi(\mathrm{d}x) \propto e^{-V(x)}\,\mathrm{d}x$, Kook and Vempala's exponential lift at Gaussian precision β has density

$$P_\beta(\mathrm{d}x\,\mathrm{d}t) \propto e^{-\beta\|x\|^2/2-dt}\mathbf{1}_{\{V(x)\le dt\}}\,\mathrm{d}x\,\mathrm{d}t.$$

Its x-marginal is π_β, and conditional on x the t-law is independent of β. The cold-start construction restricts the epigraph to one fixed bounded convex cylinder A. Let $p_\beta = P_\beta(A)$ and let ν_β be the x-marginal of $P_\beta(\cdot \mid A)$. The decisive observation is

$$p_\beta \ge \frac{1}{2} \quad (\beta \ge 0), \qquad \operatorname{Var}_{\nu_\beta}\|X\|^2 \le 2\operatorname{Var}_{\pi_\beta}\|X\|^2. \tag{9}$$

To see the first inequality, put $W_\beta(x) = V(x) + \beta\|x\|^2/2$ and $s = t + \beta\|x\|^2/(2d)$. The triangular change of variables $(x,t) \mapsto (x,s)$ sends P_β to the standard exponential lift of W_β. Kook and Vempala's same convex-truncation lemma applies to W_β: energy monotonicity decreases its radial second moment, $W_\beta(0) = V(0)$ leaves the upper lift cutoff unchanged, and $t \le s$. Thus a set of mass at least one half for the transformed lift lies inside A. The second inequality follows from $\mathrm{d}\nu_\beta/\mathrm{d}\pi_\beta \le p_\beta^{-1} \le 2$ by centering the squared radius at its π_β mean.

Consequently the Poincaré heat bound for the exact isotropic target controls the whole implemented truncated path. Replacing only the third phase of Kook and Vempala's warm-start construction by the additive precision schedule costs $\widetilde{O}(d^2\sqrt{QH_*})$. At precision zero the truncation is still 2-warm for the full lift, so their balanced full-lift sampler finishes the algorithm in $\widetilde{O}(Qd^2)$ queries.

Organization. Section 2 fixes conventions and records the finite-Rényi sampler inputs. Section 3 proves the exact overlap identity and the heat-capacity-parameterized algorithm. Section 4 proves the Poincaré and quadratic-form heat bounds. Section 5 gives the cold-start corollaries and comparisons. Section 6 proves the uniform-truncation lemma and the evaluation-oracle theorem.

2 Preliminaries and sampler inputs

All probability measures are full-dimensional unless explicitly stated otherwise. A probability measure on $\mathbb{R}^d$ is logconcave if it has a density e^{-V} on its affine support for a convex extended-real function V. A logconcave law is *isotropic* if it has mean zero and covariance I.

For a probability measure ν with finite Poincaré constant, write $\mathsf{C}_{\mathrm{P}}(\nu)$ for the least C such that

$$\operatorname{Var}_\nu f \le C\mathbb{E}_\nu \|\nabla f\|^2 \tag{10}$$

for every locally Lipschitz $f \in L^2(\nu)$. Testing linear functions shows that the worst isotropic constant $\mathcal{C}_d$ defined in Theorem 1.2 is at least one. The KLS conjecture asks whether $\sup_d \mathcal{C}_d < \infty$. We only use the proved bound $\mathcal{C}_d = O(\log(ed))$ from Klartag's isoperimetric theorem and the standard equivalence between the Cheeger and Poincaré constants for logconcave measures [3, 4, 11, 19].

For $a > 1$ and probability measures $\mu \ll \nu$, the order-a Rényi divergence is

$$\mathcal{R}_a(\mu\|\nu) = \frac{1}{a-1}\log\int\left(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\right)^a \mathrm{d}\nu. \tag{11}$$

We use the continuous extension $\mathcal{R}_1(\mu\|\nu) = \mathrm{KL}(\mu\|\nu)$ at order one. It is nondecreasing in a and obeys data processing. The associated warmness is

$$M_a(\mu\|\nu) = \|\,\mathrm{d}\mu/\,\mathrm{d}\nu\|_{L^a(\nu)} = \exp\!\left(\frac{a-1}{a}\mathcal{R}_a(\mu\|\nu)\right).$$

We use the following weak triangle inequality [15, 32].

Lemma 2.1 (Weak Rényi triangle inequality)**.** *For $a > 1$, $\theta \in (0,1)$, and probability measures μ, ν, ρ,*

$$\mathcal{R}_a(\mu\|\rho) \le \frac{a-\theta}{a-1}\mathcal{R}_{a/\theta}(\mu\|\nu) + \mathcal{R}_{(a-\theta)/(1-\theta)}(\nu\|\rho). \tag{12}$$

In particular, with $\theta = a/(2a-1)$, both orders on the right are at most $2a$, and the first coefficient is at most two.

The choice of θ in the last sentence is important. It is the order-dependent choice that justifies relaying a $2a$ guarantee through arbitrarily many phases; a fixed choice such as $2/3$ does not give the displayed pair of orders for every a.

We next state only the sampler consequences used later. They follow from the capped restricted-Gaussian sampler and the balanced uniform sampler of Kook and Vempala [15, Theorems 1.1 and 1.4 and Proposition 4.1]. We distinguish the ideal uncapped chain, the capped implementation, and the restart wrapper, since conditioning these laws interchangeably would lose the final accuracy. The exact logarithmic factors and rejection thresholds are retained in the cited source.

Theorem 2.2 (Finite-order sampler inputs). *Let $B(x_0, r) \subseteq K$ and let*

$$\mu_\beta(\,\mathrm{d}x) \propto \mathbf{1}_K(x) e^{-\beta\|x-x_0\|^2/2}\,\mathrm{d}x, \qquad 0 < \beta \le d/r^2.$$

Fix target error $\zeta \in (0, 1/10)$ and cap probability $\eta \in (0, 1/2)$. There are universal constants with the following properties.

(i) Suppose an initial law has $M_Q \le 10$ with respect to μ_β, where $Q \ge 2$. After a chosen horizon N_G, the ideal uncapped restricted-Gaussian chain has $\mathcal{R}_{2Q}$ error at most ζ. Its capped implementation agrees with that chain unless a cap is hit, has cap-hit probability at most η, and has unconditional expected membership-query cost

$$\widetilde{O}\left(\frac{d^2}{\beta r^2}\right). \tag{13}$$

It is enough that $Q \ge C \log(N_G/\eta)$. No polynomial factor in Q is suppressed in (13).

(ii) Let π be uniform on K and $\Lambda = \|\mathrm{Cov}(\pi)\|_{\mathrm{op}}$. From an initial law with $M_Q \le 10$, the ideal uncapped balanced uniform chain has $\mathcal{R}_Q$ error at most ζ after a chosen horizon N_U. The capped implementation agrees with it off an event of probability at most η and has unconditional expected cost

$$\widetilde{O}\left(\frac{Q d^2 \Lambda}{r^2}\right) \tag{14}$$

membership queries, provided $Q \ge C \log(N_U/\eta)$. Here the native bound contains a factor $\log^6(1/\zeta)$; the remaining suppressed factors are logarithmic.

The horizons are the actual iteration counts selected by the algorithms; each is bounded by its displayed query bound. A restart wrapper must draw a fresh input on every attempt. If S is the event that a coupled capped run never hits a cap, its returned law is the ideal output conditioned on S.

After the rescaling $y = (x - x_0)/r$, the variance of the restricted Gaussian is $1/(\beta r^2)$, which explains (13). The range $\beta \le d/r^2$ ensures that this variance is at least $1/d$, the initial scale used below.

3 Annealing controlled by radial heat capacity

Let π be any probability measure for which the following normalizing constants are finite. Put

$$T(x) = \|x - x_0\|^2, \qquad Z(\beta) = \int e^{-\beta T(x)/2} \pi(\,\mathrm{d}x), \qquad A(\beta) = \log Z(\beta),$$

and let π_β be the normalized tilted law. For a uniform law on a compact body, these quantities are analytic for every $\beta \ge 0$. For a general logconcave law, the same identities follow by truncation and dominated convergence; logconcavity gives all polynomial moments.

Lemma 3.1 (Heat-capacity identities). *For every $\beta \ge 0$ at which the derivatives are taken,*

$$A'(\beta) = -\frac{1}{2}\mathbb{E}_{\pi_\beta} T, \qquad A''(\beta) = \frac{1}{4}\operatorname{Var}_{\pi_\beta} T = \frac{1}{4} H(\beta). \tag{15}$$

If $\beta \ge \beta' \ge 0$ and $a > 1$, then

$$\mathcal{R}_a(\pi_\beta \| \pi_{\beta'}) = \frac{A(\beta' + a(\beta - \beta')) + (a-1)A(\beta') - aA(\beta)}{a-1}. \tag{16}$$

Consequently,

$$\mathcal{R}_a(\pi_\beta \| \pi_{\beta'}) \le \frac{a}{8}(\beta - \beta')^2 \sup_{\theta \in [\beta', \beta' + a(\beta - \beta')]} H(\theta). \tag{17}$$

The same statement with $\beta' = 0$ compares a Gaussian tilt with the untilted target.

Proof. Differentiating under the integral gives (15). If p_γ denotes the density of π_γ with respect to π, then

$$\int p_\beta^a p_{\beta'}^{1-a} \, \mathrm{d}\pi = \exp\big(A(\beta' + a(\beta - \beta')) - aA(\beta) + (a-1)A(\beta')\big),$$

which proves (16). The numerator in (16) vanishes for affine A. Applying the integral remainder formula to its second derivative, and using $A'' = H/4$, bounds it by $a(a-1)(\beta - \beta')^2 \sup H/8$. Division by $a - 1$ proves (17). $\square$

The direction in (16) is the one needed by cooling: the old, more concentrated law is measured relative to the new, less concentrated target. Reversing the divergence probes the parameter $a\beta' - (a-1)\beta$, which may be negative and need not even define a normalizable tilt for an unbounded base law.

3.1 The additive schedule

We now prove the formal version of Theorem 1.1. The theorem is stated with an internal order Q because small requested orders must be raised to meet the logarithmic rejection-cap threshold. This is a polylogarithmic issue, not a hidden polynomial factor.

Theorem 3.2 (Heat-capacity annealing)**.** *Let $K \subset \mathbb{R}^d$ be a convex body, $B(x_0, r) \subseteq K$, and let π be uniform on K. Suppose the algorithm is supplied with valid numerical values x_0, r and bounds $\overline{H}$ and $\overline{\Lambda}$ such that*

$$H(\beta) = \operatorname{Var}_{\pi_\beta} \|X - x_0\|^2 \le \overline{H} \quad (\beta \ge 0), \qquad \|\operatorname{Cov}(\pi)\|_{\mathrm{op}} \le \overline{\Lambda}. \tag{18}$$

Fix $q \ge 1$ and $\varepsilon \in (0, 1/100)$. Define the dimensionless quantity

$$\Xi = e^2 + d + q + \frac{\sqrt{\overline{H}}}{r^2} + \frac{\overline{\Lambda}}{r^2} + \frac{1}{\varepsilon}.$$

For a sufficiently large universal constant C_0, take

$$Q \ge \max\{q, 2, C_0 \log^2 \Xi\}. \tag{19}$$

This explicit choice dominates the logarithmic iteration-and-cap thresholds in Theorem 2.2; in particular, $Q = \max\{q, \widetilde{O}(1)\}$. There is a restart-until-success algorithm whose output Y satisfies

$$\mathcal{R}_q(\operatorname{law}(Y) \| \pi) \le \varepsilon$$

and whose expected membership-query complexity is

$$\widetilde{O}\left(d + \frac{d^2\sqrt{Q\overline{H}}}{r^2} + \frac{Qd^2\overline{\Lambda}}{r^2}\right). \tag{20}$$

The algorithm terminates almost surely. A finite-attempt version may instead declare failure with any prescribed probability $\delta > 0$, at the cost of an additional logarithmic factor in $1/\delta$.

Proof. We first describe an ideal uncapped pipeline and then couple all caps to it. Set

$$\beta_0 = \frac{d}{r^2}, \qquad \Delta = \min\left\{\beta_0, \frac{1}{4\sqrt{Q\overline{H}}}\right\}, \qquad L = \left\lceil \frac{\beta_0}{\Delta} \right\rceil, \qquad h = \frac{\beta_0}{L}.$$

Because $\Delta \le \beta_0$, we have

$$\frac{\Delta}{2} \le h \le \Delta, \qquad L \le 2 + \frac{4d\sqrt{Q\overline{H}}}{r^2}. \tag{21}$$

Use the positive levels

$$\beta_i = (L-i)h, \qquad i = 0, \dots, L-1,$$

followed by the untilted target at zero. Thus the smallest positive precision is h, rather than an uncontrolled residual below the mesh.

Draw exactly from π_{β_0} by rejection sampling from $N(x_0, r^2 I/d)$. The acceptance probability is at least $\mathbb{P}(\chi_d^2 \le d) > 1/2$, because $B(x_0, r) \subseteq K$; hence the expected initialization cost is universal.

For consecutive exact targets, Lemma 3.1 at order $2Q$ gives

$$\mathcal{R}_{2Q}(\pi_{\beta_i} \| \pi_{\beta_{i+1}}) \le \frac{2Q}{8} h^2 \overline{H} \le \frac{1}{64}. \tag{22}$$

Since $\beta_{L-1} = h \le \Delta$, the same calculation with $\beta' = 0$ gives

$$\mathcal{R}_{2Q}(\pi_{\beta_{L-1}} \| \pi) \le \frac{1}{64}. \tag{23}$$

Let $\widehat{\pi}_0 = \pi_{\beta_0}$. Inductively, run the ideal uncapped restricted-Gaussian chain toward $\pi_{\beta_{i+1}}$ long enough that its output $\widehat{\pi}_{i+1}$ satisfies $\mathcal{R}_{2Q}(\widehat{\pi}_{i+1} \| \pi_{\beta_{i+1}}) \le c_0$, where c_0 is a sufficiently small universal constant. Indeed, if the invariant holds at level i, Lemma 2.1 with $a = Q$ and $\theta = Q/(2Q-1)$, monotonicity, and (22) give

$$\mathcal{R}_Q(\widehat{\pi}_i \| \pi_{\beta_{i+1}}) \le 2c_0 + \frac{1}{64} < 1.$$

Thus the next call starts with $M_Q < e < 10$, and Theorem 2.2(i) restores the order-$2Q$ invariant. Equation (23) gives the same constant warmness with respect to π. Starting from $\widehat{\pi}_{L-1}$, run the ideal uncapped balanced uniform chain to an output law ρ satisfying

$$\mathcal{R}_Q(\rho \| \pi) \le \frac{\varepsilon}{2}. \tag{24}$$

We next sum the unconditional capped-call costs that will bound one full attempt. Reversing the positive grid, the Gaussian phases have precisions kh, $k = 1, \dots, L-1$. Therefore

$$\begin{aligned} \sum_{k=1}^{L-1} \widetilde{O}\left(\frac{d^2}{khr^2}\right) &\le \widetilde{O}\left(\frac{d^2}{hr^2}\right) \\ &\le \widetilde{O}\left(d + \frac{d^2\sqrt{Q\overline{H}}}{r^2}\right). \end{aligned} \tag{25}$$

Here we used (21) and $1/\Delta = \max\{r^2/d, 4\sqrt{Q\overline{H}}\}$. The uniform phase adds $\widetilde{O}(Qd^2\overline{\Lambda}/r^2)$.

Finally, put caps back into the coupled pipeline. Set $\eta = \varepsilon/8$, give every one of the at most $L-1$ Gaussian phases cap budget $\eta/(2L)$, and give the uniform phase budget $\eta/2$. The actual horizons

N_G and N_U are at most the corresponding displayed query bounds in Theorem 2.2. Using (21), their logarithms, together with $\log(L/\eta)$, are bounded by a universal constant times $\log \Xi + \log Q$. Thus the explicit choice (19), for sufficiently large C_0, meets every condition $Q \geq C \log(N/\eta')$ in that theorem. A union bound shows that the event S of no cap hit in a full attempt has probability at least $1 - \eta$.

On a cap hit, discard the attempt—including any completed annealing phases—and restart from a fresh exact sample at β_0. This also supplies the fresh warm input required when the final uniform chain is restarted. The unconditional cost of one capped attempt is at most the sum of the displayed capped-call costs, since a phase is run only when every preceding cap has succeeded. The expected number of attempts is at most $(1-\eta)^{-1}$.

It remains to check the law, rather than only the cost. Let ν be the output law of a successful attempt. The capped and ideal pipelines agree on S, so

$$\frac{\mathrm{d}\nu}{\mathrm{d}\rho} \leq \frac{1}{\mathbb{P}(S)} \leq \frac{1}{1-\eta}.$$

Directly from the definition of Rényi divergence and (24),

$$\mathcal{R}_Q(\nu \| \pi) \leq \frac{\varepsilon}{2} + \frac{Q}{Q-1} \log \frac{1}{1-\eta} \leq \varepsilon,$$

where the last inequality uses $Q \geq 2$, $\eta = \varepsilon/8$, and $\varepsilon < 1/100$. Independent full attempts therefore terminate almost surely and return exactly the law ν; monotonicity gives the desired order-q statement. Stopping after a logarithmic number of attempts gives the asserted finite-attempt variant. $\square$

Remark 3.3 (Total variation). By Pinsker's inequality, $2\mathsf{TV}(\mu,\pi)^2 \leq \mathcal{R}_1(\mu\|\pi) \leq \mathcal{R}_Q(\mu\|\pi)$. To obtain total-variation error ε, run Theorem 3.2 with Rényi error $2\varepsilon^2$. This changes only logarithmic accuracy factors.

4 Uniform radial heat capacity

This section proves Theorem 1.2. We work first with a centered isotropic logconcave probability measure π on $\mathbb{R}^d$ and write

$$\pi_\beta(\,\mathrm{d}x) \propto e^{-\beta\|x\|^2/2}\pi(\,\mathrm{d}x).$$

Let $m = m_\beta$ and $\Sigma = \Sigma_\beta$ denote its mean and covariance. Dependence on β is suppressed locally. Define

$$G_\beta = \operatorname{tr}(\Sigma_\beta^2) + m_\beta^\top \Sigma_\beta m_\beta, \qquad \Phi_\beta = -\log\det\Sigma_\beta. \tag{26}$$

Logconcavity gives all polynomial moments. Consequently the moment maps used below are right-differentiable at $\beta = 0$ and continuously differentiable for $\beta > 0$; in particular, the first-contact arguments below apply to Φ_β.

4.1 Energy, determinant, and covariance identities

Lemma 4.1 (Energy monotonicity). *For every $\beta \geq 0$,*

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}_{\pi_\beta}\|X\|^2 = -\frac{1}{2}H(\beta), \qquad \operatorname{tr}\Sigma_\beta + \|m_\beta\|^2 \leq d. \tag{27}$$

Proof. The derivative is the standard exponential-family covariance identity, or follows from Lemma 3.1. At $\beta = 0$, centered isotropy gives $\mathbb{E}_\pi \|X\|^2 = d$; the derivative is nonpositive. $\square$

Lemma 4.2 (Brascamp–Lieb and the deterministic potential bound)**.** *For $\beta > 0$,*

$$\Sigma_\beta \preceq \beta^{-1} I, \tag{28}$$

and

$$\Phi_\beta \geq d - \operatorname{tr} \Sigma_\beta \geq \|m_\beta\|^2, \qquad G_\beta \leq d + \frac{3\Phi_\beta}{\beta}. \tag{29}$$

Proof. The potential defining π_β is β-strongly convex, allowing an extended-real convex part. The Brascamp–Lieb variance inequality therefore gives (28) [2]; an indicator of a convex support follows by standard approximation.

Let $\lambda_1, \dots, \lambda_d$ be the eigenvalues of Σ_β and set $h(t) = t - 1 - \log t$. Since $h \geq 0$,

$$\sum_i h(\lambda_i) = \operatorname{tr} \Sigma_\beta - d + \Phi_\beta \geq 0.$$

Together with Lemma 4.1, this proves $\Phi_\beta \geq d - \operatorname{tr} \Sigma_\beta \geq \|m_\beta\|^2$ and also $\sum_i h(\lambda_i) \leq \Phi_\beta$.

For every $t > 0$,

$$(t-1)^2 \leq 2 \max\{1, t\}\, h(t). \tag{30}$$

If $0 < \beta \leq 1$, then $\max\{1, \lambda_i\} \leq 1/\beta$, and

$$\begin{aligned} \operatorname{tr}(\Sigma_\beta^2) &= d + 2(\operatorname{tr} \Sigma_\beta - d) + \sum_i (\lambda_i - 1)^2 \\ &\leq d + \frac{2\Phi_\beta}{\beta}. \end{aligned}$$

If $\beta \geq 1$, then $\operatorname{tr}(\Sigma_\beta^2) \leq \beta^{-1} \operatorname{tr} \Sigma_\beta \leq d$, so the same displayed upper bound remains true. Finally,

$$m_\beta^\top \Sigma_\beta m_\beta \leq \frac{\|m_\beta\|^2}{\beta} \leq \frac{\Phi_\beta}{\beta}.$$

Adding the two estimates proves (29). $\square$

Lemma 4.3 (Determinant derivative)**.** *For $\beta \geq 0$, with the identity at zero interpreted as a right derivative,*

$$\Sigma_\beta' = -\frac{1}{2} \operatorname{Cov}_{\pi_\beta}\left((X - m_\beta)(X - m_\beta)^\top, \|X\|^2 \right), \tag{31}$$

and

$$\Phi_\beta' = \frac{1}{2} \operatorname{Cov}_{\pi_\beta}(S_\beta, T), \quad S_\beta = (X - m_\beta)^\top \Sigma_\beta^{-1} (X - m_\beta), \quad T = \|X\|^2. \tag{32}$$

Proof. For an integrable statistic F that does not depend on β,

$$\frac{\mathrm{d}}{\mathrm{d}\beta} \mathbb{E}_{\pi_\beta} F = -\frac{1}{2} \operatorname{Cov}_{\pi_\beta}(F, T).$$

Differentiate the centered covariance. The two terms produced by m_β' vanish because $\mathbb{E}_{\pi_\beta}(X - m_\beta) = 0$, which proves (31). The identity $\Phi_\beta' = -\operatorname{tr}(\Sigma_\beta^{-1} \Sigma_\beta')$ then gives (32). $\square$

No monotonicity of Φ_β is asserted or needed. The argument below controls its absolute derivative.

4.2 A Poincaré proof sufficient for the query exponent

Theorem 4.4 (Poincaré heat bound)**.** *For every centered isotropic logconcave π on $\mathbb{R}^d$ and every $\beta \geq 0$,*

$$\Phi_\beta \leq 16\mathcal{C}_d^2 d\beta, \qquad G_\beta \leq 49\mathcal{C}_d^2 d, \qquad H(\beta) \leq 196\mathcal{C}_d^3 d. \tag{33}$$

Proof. Let $Y = \Sigma_\beta^{-1/2}(X - m_\beta)$. Its law is isotropic and logconcave. As a function of Y,

$$T = \left\| m_\beta + \Sigma_\beta^{1/2} Y \right\|^2, \qquad \mathbb{E}\left\|\nabla_Y T\right\|^2 = 4\left(\operatorname{tr}(\Sigma_\beta^2) + m_\beta^\top \Sigma_\beta m_\beta\right) = 4G_\beta.$$

The Poincaré inequality therefore gives

$$H(\beta) \leq 4\mathcal{C}_d G_\beta. \tag{34}$$

Likewise, $S_\beta = \|Y\|^2$, so

$$\operatorname{Var}(S_\beta) \leq 4\mathcal{C}_d d. \tag{35}$$

By Cauchy–Schwarz in (32),

$$\left|\Phi'_\beta\right| \leq 2\mathcal{C}_d \sqrt{dG_\beta}. \tag{36}$$

We claim that $\Phi_\beta \leq 16\mathcal{C}_d^2 d\beta$. At zero, $\Phi_0 = 0$, $G_0 = d$, and (34) gives $H(0) \leq 4\mathcal{C}_d d$. Moreover, $S_0 = T = \|X\|^2$ because $m_0 = 0$ and $\Sigma_0 = I$, so $\Phi'_0 = H(0)/2 \leq 2\mathcal{C}_d d$, which is strictly below the proposed barrier slope because $\mathcal{C}_d \geq 1$. If there were a first positive contact with the barrier, then Lemma 4.2 and (36) would give

$$\Phi'_\beta \leq 2\mathcal{C}_d d\sqrt{1 + 48\mathcal{C}_d^2} \leq 14\mathcal{C}_d^2 d,$$

whereas first contact from below requires $\Phi'_\beta \geq 16\mathcal{C}_d^2 d$. This contradiction proves the first estimate. Substitution into (29) gives $G_\beta \leq 49\mathcal{C}_d^2 d$, and (34) gives the last estimate. □

Combining Theorem 4.4 with $\mathcal{C}_d = O(\log(ed))$ proves $H_* = O(d\log^3(ed))$. This is already enough for $\widetilde{O}(d^{2.5})$ sampling.

The following subsection improves the heat bound itself, but is not needed for either sampling exponent.

4.3 Optional sharpening from quadratic-form variance

The input is the following recent preprint theorem. We state its status because the preceding subsection deliberately avoids depending on it for the query exponent.

Theorem 4.5 (Letwin's quadratic-form inequality)**.** *The following is Theorem 1.2 of Letwin [21]. If Y is isotropic and logconcave on $\mathbb{R}^d$, then for every symmetric matrix M,*

$$\operatorname{Var}(Y^\top MY) \leq 2\mathbb{E}\left\|\nabla(Y^\top MY)\right\|^2 = 8\operatorname{tr}(M^2). \tag{37}$$

The theorem applies without smoothness or bounded-support assumptions, and without an invertibility assumption on M.

Theorem 4.6 (Optimal-order radial heat bound)**.** *Under the hypotheses of Theorem 4.4,*

$$\Phi_\beta \leq 128d\beta, \qquad G_\beta \leq 385d, \qquad H(\beta) \leq 6160d \qquad (\beta \geq 0). \tag{38}$$

Proof. Again let $Y = \Sigma_\beta^{-1/2}(X - m_\beta)$. Decompose

$$T = \|m_\beta\|^2 + Y^\top \Sigma_\beta Y + 2(\Sigma_\beta^{1/2} m_\beta)^\top Y. \tag{39}$$

Theorem 4.5, the exact variance of the linear term, and $\operatorname{Var}(U + V) \le 2 \operatorname{Var} U + 2 \operatorname{Var} V$ yield

$$H(\beta) \le 16 \operatorname{tr}(\Sigma_\beta^2) + 8 m_\beta^\top \Sigma_\beta m_\beta \le 16 G_\beta. \tag{40}$$

Applying Theorem 4.5 to $S_\beta = Y^\top I Y$ gives $\operatorname{Var}(S_\beta) \le 8d$. Hence

$$\left|\Phi'_\beta\right| \le \frac{1}{2}\sqrt{8d \cdot 16 G_\beta} = 4\sqrt{2dG_\beta}. \tag{41}$$

Consider the barrier $128d\beta$. At zero, $G_0 = d$ and Theorem 4.5 gives $H(0) \le 8d$. As above, $S_0 = T$, so $\Phi'_0 = H(0)/2 \le 4d$. At a hypothetical first positive contact, Lemma 4.2 would give $G_\beta \le 385d$, and (41) would give

$$\Phi'_\beta \le 4\sqrt{770}\, d < 112d < 128d,$$

contradicting first contact. Thus $\Phi_\beta \le 128d\beta$. The remaining two estimates follow from (29) and (40). □

The order d cannot be improved. If π is uniform on $[-\sqrt{3}, \sqrt{3}]^d$, then it is isotropic and

$$H(0) = d \operatorname{Var}(U^2) = \frac{4}{5} d, \qquad U \sim \operatorname{Unif}[-\sqrt{3}, \sqrt{3}].$$

Corollary 4.7 (Scalar covariance)**.** *Let π be logconcave with mean x_0 and $\operatorname{Cov}(\pi) = \lambda I$. Along the spherical path (1),*

$$\sup_{\beta \ge 0} \operatorname{Var}_{\pi_\beta} \|X - x_0\|^2 \le C d \lambda^2. \tag{42}$$

Without Theorem 4.5, the right side may be replaced by $C\mathcal{C}_d^3 d\lambda^2$.

Proof. Translate by x_0 and scale by $\lambda^{-1/2}$. The transformed base law is isotropic, the precision changes from β to $\lambda\beta$, and the squared-radius statistic changes by the factor λ. Its variance therefore changes by λ^2. □

5 Cold-start consequences and comparisons

Theorem 5.1 (Cold-start sampling with scalar covariance)**.** *Let $K \subset \mathbb{R}^d$ be given by a membership oracle and let $\pi = \operatorname{Unif}(K)$. Suppose*

$$B(x_0, r) \subseteq K, \qquad x_0 = \mathbb{E}_\pi X, \qquad \operatorname{Cov}(\pi) = \lambda I.$$

Assume that x_0, r, and λ are supplied to the algorithm. For $q \ge 1$, $\varepsilon \in (0, 1/100)$, and an internal order $Q = \max\{q, \widetilde{O}(1)\}$ meeting (19), there is a restart-until-success algorithm returning Y with $\mathcal{R}_q(\operatorname{law}(Y)\|\pi) \le \varepsilon$ after

$$\widetilde{O}\left(\frac{\lambda}{r^2}(d^{5/2}\sqrt{Q} + Qd^2)\right) \tag{43}$$

expected membership queries. This statement uses only Klartag's established $\mathcal{C}_d = O(\log(ed))$ bound. Letwin's theorem merely removes the hidden factor $\mathcal{C}_d^{3/2}$ from the annealing term.

Proof. The Poincaré part of Corollary 4.7 supplies the known bound

$$\overline{H} = 196\mathcal{C}_d^3 d\lambda^2 = O\Big(d\lambda^2 \log^3(ed)\Big), \qquad \overline{\Lambda} = \lambda.$$

Substitution into Theorem 3.2 gives the displayed two terms, up to the factor $\mathcal{C}_d^{3/2}$ absorbed by $\widetilde{O}$, as well as an additive d.

To absorb the latter, translate x_0 to zero and let $\rho_K(\theta) = \sup\{t \geq 0 : t\theta \in K\}$ be the radial function of K. Since $B(0,r) \subseteq K$, $\rho_K(\theta) \geq r$ for every $\theta \in S^{d-1}$. Polar integration gives

$$d\lambda = \mathbb{E}_\pi \|X\|^2 = \frac{d}{d+2} \frac{\int_{S^{d-1}} \rho_K(\theta)^{d+2} \, \mathrm{d}\theta}{\int_{S^{d-1}} \rho_K(\theta)^d \, \mathrm{d}\theta} \geq \frac{d}{d+2} r^2.$$

Hence $\lambda/r^2 \geq 1/(d+2)$, and, since $Q \geq 2$, the term $Qd^2\lambda/r^2$ dominates d up to a universal constant. This proves (43) without invoking Theorem 4.5. Using that theorem instead sets $\overline{H} = 6160 d\lambda^2$ and only removes the indicated polylogarithmic loss. □

For $r = \lambda = 1$ and $Q = \widetilde{O}(1)$, Theorem 5.1 gives $\widetilde{O}(d^{2.5})$ expected membership queries. In the same membership-oracle model and finite-order Rényi convention, Kook and Vempala [15, Corollary 1.6] give

$$\widetilde{O}\Big(qd^2 R^{3/2} \Lambda^{1/4} + qd^2\Lambda\Big).$$

For a centered isotropic body, $R = \sqrt{d}$ and $\Lambda = 1$, making the warm-start generation term $\widetilde{O}(qd^{2.75})$. On the exact-centered scalar-covariance subclass considered here, our gain is confined to this warm-start generation term; unlike their general theorem, ours assumes that the annealing center is the exact mean and the covariance is scalar. The final $\widetilde{O}(Qd^2)$ warm sampling phase is imported from their work. Their Section 6 and Appendix A already obtain a $d^{2.5}$ exponent for certain structured families—including cubes, simplices, and bodies of revolution—under operator-norm quadratic-tilt stability. Theorem 5.1 instead applies to every centered isotropic convex body satisfying the stated inscribed-ball normalization.

The operator-stability counterexample in Kook and Vempala [15, Proposition 6.3] is compatible with our theorem. It shows that a single eigenvalue of Σ_β can grow polynomially. Our determinant argument controls $\operatorname{tr}(\Sigma_\beta^2) + m_\beta^\top \Sigma_\beta m_\beta$ and hence the variance of the scalar radial statistic; it does not claim a uniform operator-norm bound.

Limitations. The center x_0 is assumed to be the exact mean of the target. If it is not, energy monotonicity only gives a bound relative to $\mathbb{E}_\pi \|X - x_0\|^2$, and the determinant initialization $m_0 = 0$, $\Sigma_0 = I$ is lost. Likewise, scalar covariance can be handled by a scalar rescaling, but arbitrary whitening changes a spherical Gaussian tilt into an ellipsoidal one. Establishing the fully anisotropic bound

$$\sup_\beta \operatorname{Var}_{\pi_\beta} \|X - x_0\|^2 \stackrel{?}{=} O(R^2\Lambda)$$

requires a weighted determinant or spectral potential not proved here. Finally, (43) is an expected *membership-query* bound, not merely a count of accepted or proper transitions; the rejection trials, caps, and full-attempt restarts are included.

6 Evaluation-oracle logconcave sampling

Let $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be lower semicontinuous and convex, and let

$$\pi(\,\mathrm{d}x) = Z^{-1} e^{-V(x)} \, \mathrm{d}x.$$

The structural statements in this subsection allow extended-real potentials. The algorithmic theorem below returns to finite-valued V, matching the stated scope of the imported sampler propositions. The exponential lift of Kook and Vempala [13, 15] uses the convex epigraph

$$\mathcal{K} = \{(x, t) \in \mathbb{R}^d \times \mathbb{R} : V(x) \le dt\}. \tag{44}$$

One evaluation of $V(x)$ implements one membership query to $\mathcal{K}$. For a point $x_0 \in \mathbb{R}^d$ and $\beta \ge 0$, define

$$P_\beta(\,\mathrm{d}x\,\mathrm{d}t) \propto e^{-\beta\|x-x_0\|^2/2-dt}\mathbf{1}_{\mathcal{K}}(x,t)\,\mathrm{d}x\,\mathrm{d}t. \tag{45}$$

Proposition 6.1 (Exact marginal transfer)**.** *The x-marginal of P_β is*

$$\pi_\beta(\,\mathrm{d}x) \propto e^{-V(x)-\beta\|x-x_0\|^2/2}\,\mathrm{d}x.$$

Conditionally on $X = x$,

$$T = \frac{V(x) + E}{d}, \qquad E \sim \mathrm{Exp}(1),$$

and this conditional kernel does not depend on β. Therefore, for all $a > 1$ and $\beta, \beta' \ge 0$ for which the divergences are finite,

$$\mathcal{R}_a(P_\beta \| P_{\beta'}) = \mathcal{R}_a(\pi_\beta \| \pi_{\beta'}). \tag{46}$$

Proof. Integrating (45) over $t \ge V(x)/d$ produces $d^{-1}e^{-V(x)}$ times the Gaussian tilt. The displayed conditional law follows by shifting a rate-d exponential random variable. Both joint laws are obtained from their x-marginals by applying the same Markov kernel $x \mapsto \mathrm{law}(T \mid x)$. Their Radon–Nikodym derivative is therefore a function of x alone, and direct integration in (11) proves (46). □

6.1 A fixed truncation with uniform mass

Put $V_0 = \min V$ and

$$L_{\pi,g} = \{x : V(x) - V_0 \le 10d\}.$$

Suppose $x_0 \in L_{\pi,g}$, set

$$R^2 = \mathbb{E}_\pi \|X - x_0\|^2, \qquad l = \log(4e), \qquad B = \frac{V(x_0)}{d} + 5 + 13l,$$

and consider Kook and Vempala's fixed convex truncation

$$A = \mathcal{K} \cap \{(x,t) : \|x - x_0\| \le Rl,\ t \le B\}. \tag{47}$$

After translating x and the lift coordinate, this is the lifted epigraph intersected with the cylinder $B_{Rl}(0) \times [-21, 13l - 6]$ used by their algorithm [15, Equation (5.1)]. The apparent lower cutoff is already implied by the epigraph and $V(x_0) - V_0 \le 10d$.

Lemma 6.2 (Uniform mass of the fixed truncation)**.** *For every $\beta \ge 0$,*

$$p_\beta := P_\beta(A) \ge \frac{1}{2}. \tag{48}$$

Proof. Let

$$W_\beta(x) = V(x) + \frac{\beta}{2}\|x - x_0\|^2, \qquad S = T + \frac{\beta}{2d}\|X - x_0\|^2.$$

The unit-Jacobian triangular map $(x,t) \mapsto (x,s)$ sends (45) exactly to the standard exponential lift

$$e^{-ds}\mathbf{1}_{\{W_\beta(x)\le ds\}}\,\mathrm{d}x\,\mathrm{d}s$$

of W_β. Its x-marginal is π_β. By energy monotonicity, which here follows directly from

$$\frac{\mathrm{d}}{\mathrm{d}\beta}\mathbb{E}_{\pi_\beta}\|X-x_0\|^2 = -\frac{1}{2}\operatorname{Var}_{\pi_\beta}\|X-x_0\|^2 \le 0,$$

$$R_\beta^2 := \mathbb{E}_{\pi_\beta}\|X-x_0\|^2 \le R^2. \tag{49}$$

Moreover,

$$W_\beta(x_0) - \min W_\beta \le V(x_0) - V_0 \le 10d, \qquad W_\beta(x_0) = V(x_0).$$

We now reproduce the two tail estimates behind Kook and Vempala's convex truncation. Their logconcave exponential-decay lemma [13, Lemma A.7] gives

$$\mathbb{P}_{\pi_\beta}(\|X-x_0\| > R_\beta l) \le e^{-l+1} = \frac{1}{4}. \tag{50}$$

For completeness, the mean of S admits a short direct bound. Put $m_\beta = \min W_\beta$ and $F(a) = \int_{\mathbb{R}^d} e^{-a(W_\beta(x)-m_\beta)}\,\mathrm{d}x$. If x_β minimizes W_β, convexity gives, for $a \ge 1$,

$$a\left(W_\beta\left(x_\beta + \frac{z-x_\beta}{a}\right) - m_\beta\right) \le W_\beta(z) - m_\beta.$$

With $x = x_\beta + (z-x_\beta)/a$ and $\mathrm{d}x = a^{-d}\,\mathrm{d}z$, this yields $F(a) \ge a^{-d}F(1)$. For $h > 0$, it follows that

$$\frac{F(1)-F(1+h)}{h} \le \frac{1-(1+h)^{-d}}{h}F(1).$$

The integrand on the left is $e^{-U}(1-e^{-hU})/h$ for $U = W_\beta - m_\beta \ge 0$; monotone convergence as $h \downarrow 0$ therefore gives $\mathbb{E}_{\pi_\beta}(W_\beta - m_\beta) \le d$. Since conditionally on X, $S = (W_\beta(X)+E)/d$ for $E \sim \operatorname{Exp}(1)$,

$$\mathbb{E}S \le \frac{m_\beta}{d} + 1 + \frac{1}{d} \le \frac{W_\beta(x_0)}{d} + 2. \tag{51}$$

Kook and Vempala prove $\operatorname{Var}(S) \le 160$ for every such exponential lift and record the needed exponential-decay lemma [13, Lemmas 2.5 and A.7]. The S-marginal is logconcave, so applying that lemma after centering and using $\sqrt{160} < 13$ gives

$$\mathbb{P}\left(S > \frac{W_\beta(x_0)}{d} + 5 + 13l\right) \le \mathbb{P}(S - \mathbb{E}S > 13l) \le e^{-l+1} = \frac{1}{4}. \tag{52}$$

Thus a union bound assigns mass at least 1/2 to

$$\left\{\|x-x_0\| \le R_\beta l, \quad s \le \frac{W_\beta(x_0)}{d} + 5 + 13l\right\}.$$

Under the inverse triangular map this event lies in A: the radial bound follows from (49), the upper cutoff is B, and $t = s - \beta\|x-x_0\|^2/(2d) \le s$. This proves (48). □

Let $\overline{P}_\beta = P_\beta(\cdot \mid A)$ and let ν_β be its x-marginal. Because A is fixed, $\nu_\beta(\mathrm{d}x) \propto e^{-\beta\|x-x_0\|^2/2}\nu_0(\mathrm{d}x)$. The preceding lemma transfers the exact heat bound to the implemented path without requiring the truncated marginal itself to be isotropic.

Proposition 6.3 (Heat and overlap on the truncated path)**.** *For $U(x) = \|x - x_0\|^2$ and every $\beta \geq 0$,*

$$\mathrm{Var}_{\nu_\beta} U \leq 2\,\mathrm{Var}_{\pi_\beta} U. \tag{53}$$

Moreover, for $a > 1$ and $\beta, \beta' \geq 0$,

$$\mathcal{R}_a(\overline{P}_\beta \| \overline{P}_{\beta'}) = \mathcal{R}_a(\nu_\beta \| \nu_{\beta'}), \tag{54}$$

and

$$\mathcal{R}_\infty(\overline{P}_0 \| P_0) = \log \frac{1}{p_0} \leq \log 2. \tag{55}$$

Proof. Conditioning and marginalization give $\mathrm{d}\nu_\beta / \mathrm{d}\pi_\beta \leq p_\beta^{-1} \leq 2$. Hence

$$\mathrm{Var}_{\nu_\beta} U \leq \mathbb{E}_{\nu_\beta}(U - \mathbb{E}_{\pi_\beta} U)^2 \leq 2\mathbb{E}_{\pi_\beta}(U - \mathbb{E}_{\pi_\beta} U)^2,$$

which is (53). On the common fixed support A, the conditional density of T given $X = x$ is proportional to

$$e^{-dt}\mathbf{1}_{\{V(x)/d \leq t \leq B\}},$$

independently of β. The proof of Proposition 6.1 therefore gives (54). Finally, $\overline{P}_0 = P_0(\cdot \mid A)$, so (55) is immediate from Lemma 6.2. ☐

6.2 The cold-start algorithm

We now prove the formal version of Theorem 1.4. Its assumptions match the well-rounded evaluation-oracle regime of Kook and Vempala.

Theorem 6.4 (Evaluation-oracle cold start)**.** *Let $V : \mathbb{R}^d \to \mathbb{R}$ be convex and given by an evaluation oracle. Suppose $\pi(\mathrm{d}x) \propto e^{-V(x)}\,\mathrm{d}x$ is centered and isotropic and $B(0,1) \subseteq L_{\pi,g}$. Fix $q \geq 1$ and $\varepsilon \in (0, 1/100)$. For a sufficiently large universal constant C_0, take*

$$Q = \max\left\{q, 2, \left\lceil C_0 \log^2\left(e^2 + d + q + \frac{1}{\varepsilon}\right)\right\rceil\right\}. \tag{56}$$

There is a restart-until-success algorithm whose output Y satisfies

$$\mathcal{R}_q(\mathrm{law}(Y) \| \pi) \leq \varepsilon$$

and whose expected evaluation-query complexity is

$$\widetilde{O}\left(d^{5/2}\sqrt{Q} + Qd^2\right). \tag{57}$$

Here $\widetilde{O}$ suppresses logarithms in d, Q, and $1/\varepsilon$; no additional polynomial dependence on Q is hidden. No use of Theorem 4.5 is required.

Proof. We first describe an ideal uncapped attempt and verify its Rényi invariants. Rejection caps and full-attempt restarts are added at the end. Every hatted law below refers to this ideal attempt; the query costs and cap-hit probabilities refer to the capped implementations coupled to the same ideal intermediate laws.

We use the fixed set A in (47) with $x_0 = 0$. Isotropy gives $R = \sqrt{d}$, so its radial scale $D = Rl$ is $\Theta(\sqrt{d})$. In the notation of the imported construction, its annealing targets are

$$\mu_{\sigma^2,\rho}(\mathrm{d}x\,\mathrm{d}t) \propto e^{-\|x\|^2/(2\sigma^2) - \rho t}\mathbf{1}_A(x,t)\,\mathrm{d}x\,\mathrm{d}t. \tag{58}$$

When $\rho = d$ and $\beta = 1/\sigma^2$, this is exactly $\overline{P}_\beta$.

Invoke the initialization and first two ideal phases of the Rényi warm-start construction of Kook and Vempala [15, Section 5.2]. Choose their preserved base order

$$q_0 = \max\left\{2, \left\lceil C_1 \log\left(e + d + Q + \frac{1}{\varepsilon}\right)\right\rceil\right\},$$

where C_1 is a sufficiently large universal constant. Increasing C_0 if necessary ensures $q_0 \le Q$ and all rejection-cap conditions below. Their Phase I costs $\widetilde{O}(q_0^{1/2} d^{2.5})$, and their Phase II costs

$$\widetilde{O}\left(q_0^{1/2} d^{2.5} + q_0 d^2 D\right) = \widetilde{O}(d^{2.5}).$$

Cap the final inner update at $\sigma^2 = 1$; this only decreases its parameter change, so the same neighboring-target overlap estimate applies. Thus these phases produce a law $\widehat{P}_1$ satisfying

$$\mathcal{R}_{2q_0}(\widehat{P}_1 \| \overline{P}_1) \le c_0 \tag{59}$$

with $c_0 = 1/10$. At this point the lift coefficient is $\rho = d$, and the exact target is precisely $\overline{P}_1$.

By Theorem 4.4, Klartag's $\mathcal{C}_d = O(\log(ed))$ estimate, and Proposition 6.3, one may fix a universal certified constant C_H such that the truncated path has the numerical heat bound

$$\overline{H} := C_H d \log^3(ed), \qquad \operatorname{Var}_{\nu_\beta} \|X\|^2 \le \overline{H} \quad (\beta \ge 0). \tag{60}$$

Set

$$\Delta = \min\left\{1, \frac{1}{4\sqrt{Q\overline{H}}}\right\}, \qquad L = \left\lceil \frac{1}{\Delta} \right\rceil, \qquad h = \frac{1}{L},$$

and use the positive precisions $1, (L-1)h, \ldots, h$. Applying Lemma 3.1 to the base marginal of the truncated path, and then using (54), shows that consecutive exact joint targets satisfy

$$\mathcal{R}_{2q_0}(\overline{P}_{kh} \| \overline{P}_{(k-1)h}) \le \frac{2q_0}{8} h^2 \overline{H} \le \frac{1}{64} \tag{61}$$

for $k = 2, \ldots, L$.

At each new positive precision, run Kook and Vempala's truncated proximal sampler $\mathsf{PS}_{\mathrm{ann}}$ from order q_0 to order $2q_0$ and restore the invariant $\mathcal{R}_{2q_0} \le c_0$. The weak triangle inequality and (61) keep its input warmness below 10. Proposition 5.2 of Kook and Vempala [15] charges, at target $\beta = 1/\sigma^2$,

$$\widetilde{O}\left(d^2(\sigma^2 \vee 1)\right) = \widetilde{O}\left(d^2(\beta^{-1} \vee 1)\right) \tag{62}$$

evaluation queries. Consequently all positive-precision calls cost

$$\sum_{k=1}^{L-1} \widetilde{O}\left(\frac{d^2}{kh}\right) = \widetilde{O}\left(\frac{d^2}{h}\right) = \widetilde{O}\left(d^2 \sqrt{Q\overline{H}}\right). \tag{63}$$

At the last positive target $\overline{P}_h$, make one additional same-target call to $\mathsf{PS}_{\mathrm{ann}}$, now from order q_0 to order $2Q$, obtaining a law $\widehat{P}_h$ with $\mathcal{R}_{2Q}(\widehat{P}_h \| \overline{P}_h) \le c_0$. The target-order dependence of this call is only logarithmic, and its cost is absorbed by (63). The terminal exact overlap is

$$\mathcal{R}_{2Q}(\overline{P}_h \| \overline{P}_0) \le \frac{2Q}{8} h^2 \overline{H} \le \frac{1}{64}.$$

The weak triangle inequality therefore gives

$$\mathcal{R}_Q(\widehat{P}_h \| \overline{P}_0) < 1 \tag{64}$$

because $2c_0 + 1/64 < 1$.

Let ρ be the x-marginal of $\widehat{P}_h$ and let ν_0 be the x-marginal of $\overline{P}_0$. By data processing, $M_Q(\rho\|\nu_0) < e$. Also $\mathrm{d}\nu_0/\mathrm{d}\pi \le p_0^{-1} \le 2$, and hence

$$M_Q(\rho\|\pi) \le 2^{(Q-1)/Q} M_Q(\rho\|\nu_0) < 2e < 10. \tag{65}$$

Resample $T = (V(X) + E)/d$ conditionally on $X \sim \rho$. Because this is the same conditional kernel as under P_0, the resulting joint warmness relative to P_0 is exactly $M_Q(\rho\|\pi) < 10$. Now run the full, untruncated exponential-lift sampler $\mathsf{PS}_{\mathrm{exp}}$ of Kook and Vempala. Its balanced finite-order guarantee [15, Proposition 5.1], together with isotropy $\|\mathrm{Cov}(\pi)\|_{\mathrm{op}} = 1$, produces order-$Q$ error at most $\varepsilon/2$ using $\widetilde{O}(Qd^2)$ additional evaluation queries.

Put back the capped implementations coupled to the ideal attempt just analyzed. Give the calls total cap budget $\eta = \varepsilon/8$, and restart the entire pipeline from a fresh initialization if any cap is hit. The phase count and all iteration horizons have logarithms bounded by a universal multiple of $\log(e^2 + d + Q + 1/\varepsilon)$. Thus one may choose q_0 polylogarithmically and then take C_0 in (56) large enough to meet every cap condition. The unconditional expected cost of one capped attempt is bounded by the sum above, because a call is made only if all earlier calls succeeded. A union bound makes one attempt successful with probability at least $1 - \eta$. Let ρ_* be the ideal final x-law, let S be the event that no cap is hit, and let $\nu = \mathrm{law}(Y \mid S)$ be the law returned by a successful attempt. The coupling gives

$$\frac{\mathrm{d}\nu}{\mathrm{d}\rho_*} \le \frac{1}{\mathbb{P}(S)} \le \frac{1}{1-\eta}.$$

Directly from the definition of Rényi divergence,

$$\mathcal{R}_Q(\nu\|\pi) \le \mathcal{R}_Q(\rho_*\|\pi) + \frac{Q}{Q-1} \log \frac{1}{1-\eta} \le \varepsilon.$$

Here the ideal final sampler was run to error at most $\varepsilon/2$, $Q \ge 2$, $\eta = \varepsilon/8$, and $\varepsilon < 1/100$. Full restarts increase the expected cost by only a constant factor and terminate almost surely.

Finally, (60) turns (63) into $\widetilde{O}(d^{5/2}\sqrt{Q})$. Adding Phases I–II and the final balanced call proves (57); monotonicity in the Rényi order gives the requested order q. $\square$

For comparison, the evaluation-oracle bound of Kook and Vempala [15, Corollary 1.9] is

$$\widetilde{O}\Big(d^{2.5} + qd^2 R^{3/2}(\Lambda^{1/4} \vee 1) + qd^2(\Lambda \vee 1)\Big),$$

up to logarithmic accuracy factors, where $R^2 = \mathbb{E}_\pi \|X\|^2$ and $\Lambda = \|\mathrm{Cov}(\pi)\|_{\mathrm{op}}$. In centered isotropic position this is $\widetilde{O}(qd^{2.75})$. Theorem 6.4 replaces the third-phase term on exactly that subclass; it retains their ground-set normalization, preprocessing, truncated sampler, and final full-lift sampler. The later unified analysis of Kook and Vempala [14] improves the last warm-start sampling term, but does not supply the order-boosting truncated sampler used in the annealing phases. As throughout the paper, exact centering and isotropy are assumptions of the result, not outputs of a rounding procedure.

References

[1] David L. Applegate and Ravi Kannan. Sampling and integration of near log-concave functions. In *Proceedings of the Twenty-Third Annual ACM Symposium on Theory of Computing*, pages 156–163. ACM, 1991. doi: 10.1145/103418.103439.

[2] Herm Jan Brascamp and Elliott H. Lieb. On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4):366–389, 1976. doi: 10.1016/0022-1236(76)90004-5.

[3] Peter Buser. A note on the isoperimetric constant. *Annales Scientifiques de l'École Normale Supérieure*, 15(2):213–230, 1982. doi: 10.24033/asens.1426.

[4] Jeff Cheeger. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in Analysis*, pages 195–199. Princeton University Press, 1970.

[5] Yongxin Chen, Sinho Chewi, Adil Salim, and Andre Wibisono. Improved analysis for a proximal algorithm for sampling. In *Proceedings of the Thirty-Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 2984–3014. PMLR, 2022.

[6] Ben Cousins and Santosh S. Vempala. Bypassing KLS: Gaussian cooling and an $O^*(n^3)$ volume algorithm. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, pages 539–548. ACM, 2015. doi: 10.1145/2746539.2746563.

[7] Ben Cousins and Santosh S. Vempala. Gaussian cooling and $O^*(n^3)$ algorithms for volume and Gaussian volume. *SIAM Journal on Computing*, 47(3):1237–1273, 2018. doi: 10.1137/15M1054250.

[8] Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM*, 38(1):1–17, 1991. doi: 10.1145/102782.102783.

[9] Ravi Kannan, László Lovász, and Miklós Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13(3–4):541–559, 1995. doi: 10.1007/BF02574061.

[10] Ravi Kannan, László Lovász, and Miklós Simonovits. Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Structures & Algorithms*, 11(1):1–50, 1997. doi: 10.1002/(SICI)1098-2418(199708)11:1<1::AID-RSA1>3.0.CO;2-X.

[11] Bo'az Klartag. Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inveniendi Analytica*, 2023. doi: 10.15781/jsjy-0b06. Paper No. 4, 17 pp.

[12] Yunbum Kook and Santosh S. Vempala. Faster logconcave sampling from a cold start in high dimension. In *2025 IEEE 66th Annual Symposium on Foundations of Computer Science*, pages 997–1006. IEEE, 2025. doi: 10.1109/FOCS63196.2025.00052.

[13] Yunbum Kook and Santosh S. Vempala. Sampling and integration of logconcave functions by algorithmic diffusion. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 924–932. ACM, 2025. doi: 10.1145/3717823.3718202. URL https://arxiv.org/abs/2411.13462. Full version: arXiv:2411.13462v1.

[14] Yunbum Kook and Santosh S. Vempala. A unified complexity bound for logconcave sampling, 2026. URL https://arxiv.org/abs/2606.12694. arXiv:2606.12694v1.

[15] Yunbum Kook and Santosh S. Vempala. Zeroth-order logconcave sampling, 2026. URL https://arxiv.org/abs/2507.18021. arXiv:2507.18021v2.

[16] Yunbum Kook and Matthew S. Zhang. Rényi-infinity constrained sampling with d^3 membership queries. In *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 5278–5306. SIAM, 2025. doi: 10.1137/1.9781611978322.181.

[17] Yunbum Kook, Santosh S. Vempala, and Matthew S. Zhang. In-and-Out: Algorithmic diffusion for sampling convex bodies. In *Advances in Neural Information Processing Systems*, volume 37, pages 108354–108388, 2024.

[18] Yunbum Kook, Santosh S. Vempala, and Matthew S. Zhang. In-and-Out: Algorithmic diffusion for sampling convex bodies. *Random Structures & Algorithms*, 68(3):e70061, 2026. doi: 10.1002/rsa.70061.

[19] Michel Ledoux. Spectral gap, logarithmic Sobolev constant, and geometric bounds. In *Surveys in Differential Geometry IX*, volume 9, pages 219–240. International Press, 2004. doi: 10.4310/SDG.2004.v9.n1.a6.

[20] Yin Tat Lee, Ruoqi Shen, and Kevin Tian. Structured logconcave sampling with a restricted Gaussian oracle. In *Proceedings of the Thirty-Fourth Conference on Learning Theory*, volume 134 of *Proceedings of Machine Learning Research*, pages 2993–3050. PMLR, 2021.

[21] Brayden Letwin. The KLS constant is $O(\log^{1/4} n)$, 2026. URL https://arxiv.org/abs/2607.24164. arXiv:2607.24164v1.

[22] László Lovász. Hit-and-run mixes fast. *Mathematical Programming*, 86(3):443–461, 1999. doi: 10.1007/s101070050099.

[23] László Lovász and Miklós Simonovits. The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. In *31st Annual Symposium on Foundations of Computer Science*, pages 346–354. IEEE, 1990. doi: 10.1109/FSCS.1990.89543.

[24] László Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Structures & Algorithms*, 4(4):359–412, 1993. doi: 10.1002/rsa.3240040402.

[25] László Lovász and Santosh S. Vempala. Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. In *47th Annual IEEE Symposium on Foundations of Computer Science*, pages 57–68. IEEE, 2006. doi: 10.1109/FOCS.2006.28.

[26] László Lovász and Santosh S. Vempala. Hit-and-run from a corner. *SIAM Journal on Computing*, 35(4):985–1005, 2006. doi: 10.1137/S009753970544727X.

[27] László Lovász and Santosh S. Vempala. Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. *Journal of Computer and System Sciences*, 72(2):392–417, 2006. doi: 10.1016/j.jcss.2005.08.004.

[28] László Lovász and Santosh S. Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures & Algorithms*, 30(3):307–358, 2007. doi: 10.1002/rsa.20135.

[29] Peter Salamon, James D. Nulton, John R. Harland, Jacob Pedersen, George Ruppeiner, and Luby Liao. Simulated annealing with constant thermodynamic speed. *Computer Physics Communications*, 49(3):423–428, 1988. doi: 10.1016/0010-4655(88)90003-3.

[30] Robert L. Smith. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32(6):1296–1308, 1984. doi: 10.1287/opre.32.6.1296.

[31] Daniel Štefankovič, Santosh S. Vempala, and Eric Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *Journal of the ACM*, 56(3): 18:1–18:36, 2009. doi: 10.1145/1516512.1516520.

[32] Tim van Erven and Peter Harremoës. Rényi divergence and Kullback–Leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820, 2014. doi: 10.1109/TIT.2014.2320500.