Sampling and Integration of Logconcave Functions by Algorithmic Diffusion

Yunbum Kook
Georgia Institute of Technology
Atlanta, USA
yb.kook@gatech.edu

Santosh S. Vempala
Georgia Institute of Technology
Atlanta, USA
vempala@gatech.edu

Abstract

We study the complexity of sampling, rounding, and integrating arbitrary logconcave functions given an evaluation oracle. Our new approach provides the first complexity improvements in nearly two decades for general logconcave functions for all three problems, and matches the best-known complexities for the special case of uniform distributions on convex bodies. For the sampling problem, our output guarantees are significantly stronger than previously known, and lead to a streamlined analysis of statistical estimation based on dependent random samples.

CCS Concepts

• **Theory of computation** → **Random walks and Markov chains**.

Keywords

Sampling, Functional inequalities, Rounding, Integration

ACM Reference Format:
Yunbum Kook and Santosh S. Vempala. 2025. Sampling and Integration of Logconcave Functions by Algorithmic Diffusion. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing (STOC '25), June 23–27, 2025, Prague, Czechia.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3717823.3718202

1 INTRODUCTION

Sampling and integration of logconcave functions are fundamental problems with numerous applications and important special cases such as uniform distributions over convex bodies and strongly logconcave densities. The study of these problems has led to many useful techniques. Both mathematically and algorithmically, general logconcave functions typically provide the "right" general abstraction. For example, many classical inequalities for convex bodies have natural extensions to logconcave functions (e.g., Grünbaum's theorem, Brunn-Minkowski and Prékopa-Leindler, isotropic constant, etc.). The KLS hyperplane conjecture, first motivated by the analysis of the Ball Walk for sampling convex bodies, is in the setting of general logconcave densities. The current fastest algorithm for estimating the volume of a convex body crucially uses sampling from a sequence of logconcave densities, which is provably more efficient than using a sequence of uniform distributions. Sampling logconcave densities has many other applications as well, such as portfolio optimization, simulated annealing, Bayesian inference, differential privacy etc.

Sampling in high dimension is done by randomized algorithms based on Markov chains. These chains are set up to have a desired stationary distribution, which is relatively easy to ensure. For example, to sample uniformly, it suffices that the Markov chain is symmetric. Generally, to sample proportional to a desired function, it suffices to ensure the "detailed balance" condition. The main challenge is showing rapid mixing of the Markov chain, i.e., the convergence rate to the stationary distribution is bounded by a (small) polynomial in the dimension and other relevant parameters.

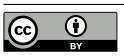
The traditional analysis of Markov chains for sampling high-dimensional distributions proceeds by analyzing the *conductance* of the Markov chain, the minimum conditional escape probability over all subsets of the state space of measure at most half. Bounding this is done by relating probabilistic (one-step distribution) distance to geometric distance, and then using a purely isoperimetric inequality for subsets of the support. Indeed, this approach led to several interesting questions and useful techniques, in particular the development of isoperimetric inequalities and the discovery of (nearly) tight bounds in many settings. To describe background and known results, let us first define the sampling problem (readers familiar with the problem can skip to §1.1).

Model and Problems. We assume access to an integrable logconcave function $\exp(-V(x))$ in $\mathbb{R}^n$, via an evaluation oracle for a convex function $V : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$. We also assume that there exists a point x_0 and parameters r, R such that for the distribution π with density $\mathrm{d}\pi \propto \exp(-V)\,\mathrm{d}x$, a ball of radius r centered at x_0 is contained in the level set of π of measure $1/8$[1] and $\mathbb{E}_\pi[\|X - x_0\|^2] \le R^2$. We refer to this as a *well-defined function oracle*[2], denoted by $\mathrm{Eval}_{\mathcal{P}}(V)$ where $\mathcal{P}$ indicates access to the actual values of parameters in $\mathcal{P}$ (e.g., $\mathrm{Eval}_{x_0,R}(V)$ presents both x_0 and R, while $\mathrm{Eval}(V)$ does not provide any parameters).

Given this oracle, in this paper[3] we consider three central problems: (1) sample from the distribution π, (2) find an affine transformation that places π in near-isotropic position, and (3) estimate the integral $\int e^{-V(x)}\,\mathrm{d}x$ (i.e., the normalization constant of π). We measure the complexity in terms of the number of oracle calls and the total number of arithmetic operations.

Complexity of Logconcave Sampling. Sampling general logconcave functions, as an algorithmic problem, was first studied by Applegate and Kannan [2], who established an algorithm with complexity polynomial in the dimension assuming Lipschitzness of the

[1] We use the standard convention of $1/8$; any constant bounded away from 1 would suffice. In fact, we will use an even weaker condition.
[2] Without loss of generality, we will assume by scaling that $r = 1$.
[3] The full paper is available on arXiv at [20].



STOC '25, Prague, Czechia

ACM ISBN 979-8-4007-1510-5/25/06
https://doi.org/10.1145/3717823.3718202

function over its support. They did this by creating a sufficiently fine grid (so that the function changed by at most a small constant within each grid cell) and used a discrete "grid walk". They sampled a grid point and used rejection sampling to output a point from the cube associated with the grid point. The output distribution is guaranteed to be within desired total variation (TV) distance from the target distribution. Following many developments, Lovász and Vempala improved the bound to scale as n^2R^2/r^2 [31, 34]. More specifically, [34, Theorem 2.2, 2.3] analyzed the Ball Walk and Hit-and-Run and proved a bound of

$$\frac{M^4}{\varepsilon^4}\frac{n^2R^2}{r^2}\log\frac{M}{\varepsilon}$$

steps/queries to reach a distribution within ε TV-distance of the target starting from an M-warm distribution. [31, Theorem 1.1] showed that Hit-and-Run achieves the same guarantee in

$$\frac{n^2R^2}{r^2}\log^{O(1)}\frac{nMR}{\varepsilon r}$$

steps. We note that in all previous work on general logconcave sampling the output guarantees are in TV-distance. It is often desirable to have stronger guarantees such as the KL or Rényi divergences.

Classical Sampling Algorithms. For target density proportional to e^{-V}, the Ball Walk with parameter δ proceeds as follows: at a point x with $e^{-V(x)} > 0$, sample a uniform random point y in the ball of radius δ centered at x; go to y with probability $\min(1, \frac{e^{-V(y)}}{e^{-V(x)}})$, staying at x with the remaining probability. Hit-and-Run does not need a parameter: at current point x, pick a uniform random line ℓ through x, and go to a random point y along ℓ with probability proportional to $e^{-V(y)}$ (marginal along ℓ). For technical reasons, both of these walks (and almost all other walks) need to be *lazy*, i.e., they do nothing with probability $1/2$, and do the above step with probability $1/2$.

Sampling and Isoperimetry. The complexity of sampling is intuitively tied closely to the isoperimetry of the target distribution. If the target has poor isoperimetry (roughly, a small measure surface can partition the support into two large measure subsets), then any "local" Markov chain will have difficulty moving from one large subset to its complement. There are two alternative views of isoperimetry — functional and geometric. We recap the definitions of isoperimetric constants.

Definition 1.1. We say that a probability measure π on $\mathbb{R}^n$ satisfies a *Poincaré inequality* (PI) with parameter $C_{\mathsf{PI}}(\pi)$ if for all smooth functions $f : \mathbb{R}^n \to \mathbb{R}$,

$$\operatorname{var}_\pi f \le C_{\mathsf{PI}}(\pi)\,\mathbb{E}_\pi[\|\nabla f\|^2], \tag{PI}$$

where $\operatorname{var}_\pi f = \mathbb{E}_\pi[|f - \mathbb{E}_\pi f|^2]$.

The Poincaré inequality is implied by the generally stronger log-Sobolev inequality.

Definition 1.2. We say that a probability measure π on $\mathbb{R}^n$ satisfies a *log-Sobolev inequality* (LSI) with parameter $C_{\mathsf{LSI}}(\pi)$ if for all smooth functions $f : \mathbb{R}^n \to \mathbb{R}$,

$$\operatorname{Ent}_\pi(f^2) \le 2C_{\mathsf{LSI}}(\pi)\,\mathbb{E}_\pi[\|\nabla f\|^2], \tag{LSI}$$

where $\operatorname{Ent}_\pi(f^2) := \mathbb{E}_\pi[f^2\log f^2] - \mathbb{E}_\pi[f^2]\log\mathbb{E}_\pi[f^2]$.

We define $\pi(\partial S) := \liminf_{\varepsilon\downarrow 0}\frac{\pi(S_\varepsilon)-\pi(S)}{\varepsilon}$, where $S_\varepsilon := \{x : d(x,S) \le \varepsilon\}$, for the geometric view.

Definition 1.3. The *Cheeger constant* $C_{\mathsf{Ch}}(\pi)$ of a probability measure π on $\mathbb{R}^n$ is defined as

$$C_{\mathsf{Ch}}(\pi) = \inf_{S\subset\mathbb{R}^n,\pi(S)\le\frac{1}{2}}\frac{\pi(\partial S)}{\pi(S)}.$$

Definition 1.4. The *log-Cheeger constant* $C_{\mathsf{logCh}}(\pi)$ of a probability measure π on $\mathbb{R}^n$ is defined as

$$C_{\mathsf{logCh}}(\pi) = \inf_{S\subset\mathbb{R}^n,\pi(S)\le\frac{1}{2}}\frac{\pi(\partial S)}{\pi(S)\sqrt{\log(1/\pi(S))}}.$$

It is known that for logconcave measures, $C_{\mathsf{PI}}(\pi) = \Theta(C_{\mathsf{Ch}}^{-2}(\pi))$ [5, 6, 25] and $C_{\mathsf{LSI}}(\pi) = \Theta(C_{\mathsf{logCh}}^{-2}(\pi))$ [24]. Bounding these constants has been a major research topic for decades. In recent years, following many improvements, it has been shown that for isotropic logconcave measures, $C_{\mathsf{PI}}(\pi) \lesssim \log n$ [16] and for isotropic logconcave ones with support of diameter D, we have $C_{\mathsf{LSI}}(\pi) \lesssim D$ [27]. The former is conjectured to be $O(1)$ (the KLS conjecture), and the latter is the best possible.

The significance of these constants for algorithmic sampling became clear with the analysis of the Ball Walk, where the bound on its convergence from a warm start to a logconcave distribution depends directly on $C_{\mathsf{Ch}}^{-2}(\pi)$ [15] (this was the original motivation for the KLS conjecture) — the mixing time of the Ball Walk starting from an M-warm distribution to reach ε TV-distance is bounded by $O(n^2C_{\mathsf{Ch}}^{-2}\operatorname{poly}\frac{M}{\varepsilon})$. For the special case of uniformly sampling convex bodies, using the notion of a Speedy Walk, this can be improved to $O(Mn^2C_{\mathsf{Ch}}^{-2}\log\frac{M}{\varepsilon})$, but using it in the analysis entails handling several technical difficulties. In principle, this might be extendable to general logconcave functions, albeit with formidable technical complications. We take a different approach.

Sampling, Isoperimetry, and Diffusion. The connection of isoperimetric constants to the convergence of continuous-time *diffusion* is a classical subject. The Langevin diffusion is a canonical stochastic differential equation (SDE) given by

$$\mathrm{d}X_t = -\nabla V(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}B_t, \quad X_0 \sim \nu_0,$$

where B_t is the standard Brownian process, and V is a smooth function. Then, under mild conditions, the law ν_t of X_t converges to the distribution with density ν proportional to e^{-V} in the limit. Moreover, it can be shown that

$$\chi^2(\nu_t \,\|\, \nu) \le \exp\Big(-\frac{2t}{C_{\mathsf{PI}}(\nu)}\Big)\,\chi^2(\nu_0 \,\|\, \nu),$$
$$\mathsf{KL}(\nu_t \,\|\, \nu) \le \exp\Big(-\frac{2t}{C_{\mathsf{LSI}}(\nu)}\Big)\,\mathsf{KL}(\nu_0 \,\|\, \nu).$$

A natural idea for a sampling algorithm is to discretize the diffusion equation. This discretized algorithm was shown to converge in $\operatorname{poly}(n,\beta,C_{\mathsf{PI}}(\nu))$ and $\operatorname{poly}(n,\beta,C_{\mathsf{LSI}}(\nu))$ iterations under (PI) [8] and (LSI) [37] respectively, assuming β-Lipschitzness of ∇V. Can this approach be extended to sampling general logconcave densities without the gradient Lipschitzness?

Algorithm 1 Proximal Sampler $\mathsf{PS}_{\exp}$

Input: initial pt. $z_0 \sim \pi_0 \in \mathcal{P}(\mathbb{R}^{n+1})$, $\mathcal{K} = \{(x,t) : V(x) \le nt\}$, $k \in \mathbb{N}$, threshold N, variance h. **Output:** $z_{k+1} = (x_{k+1}, t_{k+1})$.

1: **for** $i = 0, \dots, k$ **do**
2: Sample $y_{i+1} \sim \mathcal{N}(z_i, hI_{n+1})$.
3: Sample $z_{i+1} \sim \mathcal{N}(y_{i+1} - hne_{n+1}, hI_{n+1})|_{\mathcal{K}}$.
4: ($\uparrow$) Repeat $z_{i+1} \sim \mathcal{N}(y_{i+1} - hne_{n+1}, hI_{n+1})$ until $z_{i+1} \in \mathcal{K}$. If $\#\text{attempts}_i \ge N$, declare **Failure**.
5: **end for**

Algorithmic Diffusion. By *algorithmic diffusion*, we mean the general idea of discretizing a diffusion process and proving guarantees on the discretization error and query complexity. Recently, [21] showed that such an algorithm (called In-and-Out there), which can be viewed as a Proximal Sampler, works for sampling from the uniform distribution over a convex body and recovers state-of-the-art complexity guarantees with substantially improved output guarantees (Rényi-divergences). The guarantee was extended to the Rényi-infinity (or pointwise) distance as well [23]. A nice aspect of the analysis is that it shows convergence directly in terms of isoperimetric constants of target distributions. As a result, it provides a unifying point of view for analysis without requiring the use of technically sophisticated tools such as the Speedy Walk and s-conductance [15, 30, 34]. We discuss the approach in more detail presently.

This brings us to the main motivation of the current paper: *Can we use algorithmic diffusion to obtain faster algorithms with stronger guarantees for sampling, rounding, and integrating general logconcave functions?*

1.1 Results

In this paper, we propose new approaches for general logconcave sampling, rounding and integration, leading to the first improvements in the complexity of these problems in nearly two decades [31]. Our methods crucially rely on a reduction from general logconcave sampling to *exponential sampling* in one higher dimension. With this reduction in hand, we develop a new framework for sampling and improve the complexity of four fundamental problems: (1) logconcave sampling from an $O(1)$-warm start, (2) warm-start generation, (3) isotropic rounding and (4) integration. For each problem, our improved complexity for general logconcave distributions matches the current best complexity for the uniform distribution over a convex body, a setting that has received much attention for more than three decades [9, 11, 15, 21, 23, 28–30, 32]. We now present our main results, followed by a detailed discussion of the challenges and techniques in §1.2, and notation and definitions in §1.3.

Result 1: Logconcave Sampling from a Warm Start. When a logconcave distribution $\pi^X \propto e^{-V}$ is presented with the evaluation oracle for V, we instead consider the "lifted" distribution

$$\pi(x,t) \propto \exp(-nt) \cdot \mathbb{1}[(x,t) : V(x) \le nt]$$

in the augmented (x,t)-space, the X-marginal of which is exactly π^X. We quantify how this affects parameters pertinent to sampling (e.g., isoperimetric constants).

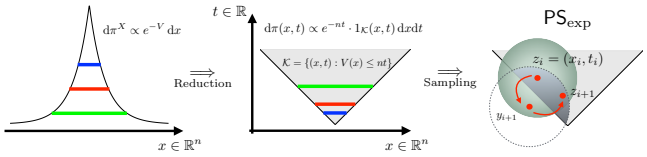


Figure 1.1: Reduction to an exponential distribution and sampling via the Proximal Sampler $\mathsf{PS}_{\exp}$.

To sample from the exponential target π using a given $\mathcal{R}_\infty$ warm-start, we use the Proximal Sampler (see Figure 1.1) with an evaluation oracle that returns nt conditioned on $V(x) \le nt$. One iteration of the sampler is reversible with respect to the stationary distribution π, while it can be interpreted as a composition of two complimentary diffusion processes. Discarding the T-component of a sample, we are left with the X-component with law close to our desired target π^X. In this regard, our sampler is a natural generalization of In-and-Out [21], the Proximal Sampler for uniform sampling. Concretely, we establish mixing guarantees and query complexities of our sampler for the target logconcave distribution π^X in terms of its Poincaré constant (which in general satisfies $\|\mathrm{Cov}\,\pi^X\| \le C_{\mathsf{PI}}(\pi^X) \lesssim \|\mathrm{Cov}\,\pi^X\| \log n$ [16]).

Theorem 1.5. *For any logconcave distribution π^X specified by a well-defined function oracle* $\mathsf{Eval}(V)$*, for any given $\eta, \varepsilon \in (0,1)$, $q \ge 2$, and π_0^X with $\mathcal{R}_\infty(\pi_0^X \,\|\, \pi^X) = \log M$, we can use the* Proximal Sampler $\mathsf{PS}_{\exp}$ *(Algorithm 1) with suitable choices of parameters, so that with probability at least $1-\eta$, we obtain a sample X such that $\mathcal{R}_q(\mathrm{law}\,X \,\|\, \pi^X) \le \varepsilon$, using $\widetilde{O}(qMn^2\,(\|\mathrm{Cov}\,\pi^X\| \vee 1)\,\mathrm{polylog}\,\frac{1}{\eta\varepsilon})$ evaluation queries in expectation.*

See the full paper for a more detailed description of the heat flow perspective and its benefits for sampling. Our specific setting of parameters can be found in the full paper. In comparison with the previous best complexity of $\widetilde{O}(n^2\,\mathrm{tr}(\mathrm{Cov}\,\pi^X)\,\mathrm{polylog}\,\frac{M}{\varepsilon})$, which is for TV-distance, the Proximal Sampler $\mathsf{PS}_{\exp}$ achieves a provably better rate (since $\|\mathrm{Cov}\,\pi^X\| \le \mathrm{tr}(\mathrm{Cov}\,\pi^X)$) from an $O(1)$-warm start, and moreover does so in general $\mathcal{R}_q$-divergences. This mixing rate matches the previously known best rate for the uniform sampling by the Ball Walk [15] as well as In-and-Out [21].

Remark 1.6 (Connection to well-conditioned settings). A recent line of research in sampling literature has studied the complexity of samplers for *well-conditioned* distributions, which refers to unconstrained α-strongly logconcave and β-smooth distributions with condition number $\kappa := \beta/\alpha$. For instance, zeroth-order samplers have query complexities $\widetilde{O}(\kappa n \log 1/\varepsilon)$ for various metrics from a warm start (actually, from a proper feasible start), including the metropolized Gaussian walk [1], Hit-and-Run [3], Proximal Sampler [7, 26]. We can show that our reduction scheme yields a complexity of $\widetilde{O}((\kappa n + n^2)\log 1/\eta\varepsilon)$ for well-conditioned distributions, revealing a connection to this line of research.

Result 2: Warm-start Generation (Sampling without a Warm Start). The Proximal Sampler $\mathsf{PS}_{\exp}$ assumes access to a warm start in $\mathcal{R}_\infty$, and generating such a good warm-start is an important and challenging algorithmic problem in its own right. We propose Tilted Gaussian Cooling which generates an $O(1)$-warm start in

the $\mathcal{R}_\infty$-divergence for any target logconcave distribution. This algorithm generalizes Gaussian Cooling [9], the known method for generating a warm-start for the uniform distribution over a convex body. The high-level idea is to follow a sequence $\{\mu_i\}_{i\in[m]}$ of distributions, where μ_1 is easy to sample from, μ_i and μ_{i+1} are close in some probability divergences, and μ_m is the desired target.

Our algorithm first reduces the original target to the exponential distribution and then follows annealing distributions of the form

$$\mu_{\sigma^2,\rho}(x,t) \propto \exp\Big(-\frac{1}{2\sigma^2}\|x\|^2 - \rho t\Big) \times \mathbb{1}\left[\{V(x) \le nt\} \cap \{\|x\| = O(R), |t| = O(1)\}\right],$$

with a carefully chosen schedule for updating the parameters σ^2 and ρ. Roughly speaking, we update the X-marginal through Gaussian Cooling and the T-marginal using an exponential *tilt*. In order to move across the annealing distributions, Tilted Gaussian Cooling requires an efficient sampler for the intermediate annealing distributions, ideally with guarantees in the $\mathcal{R}_\infty$-divergence in order to relay $\mathcal{R}_\infty$-warmness guarantees along the annealing scheme. We use the Proximal Sampler for these intermediate distributions with a $\mathcal{R}_\infty$-divergence guarantee, namely the Proximal Sampler $\mathsf{PS}_{\mathsf{ann}}$. In our computational model, this sampler started at a previous annealing distribution returns a sample with law μ satisfying $\mathcal{R}_\infty(\mu \parallel \mu_{\sigma^2,\rho}) \le \varepsilon$, using $\widetilde{O}(n^2\sigma^2 \operatorname{polylog} R/\eta\varepsilon)$ evaluation queries in expectation.

By sampling these annealing distributions through $\mathsf{PS}_{\mathsf{ann}}$, Tilted Gaussian Cooling obtains an $O(1)$-warm start for π (and hence for the desired target π^X); then runs the Proximal Sampler $\mathsf{PS}_{\mathsf{exp}}$ one final time to obtain a sample with $\mathcal{R}_\infty$-guarantees (this is a very strong notion of probability divergence that recovers all commonly used distances such as TV, KL, χ^2, or $\mathcal{R}_q$).

Theorem 1.7. *For any logconcave distributions π^X specified by $\mathsf{Eval}_{x_0,R}(V)$, for any given $\eta,\varepsilon \in (0,1)$,* Tilted Gaussian Cooling *with probability at least $1-\eta$, returns a sample with law ν such that $\mathcal{R}_\infty(\nu \parallel \pi^X) \le \varepsilon$, using $\widetilde{O}(n^2(R^2 \vee n)\operatorname{polylog} 1/\eta\varepsilon)$ evaluation queries in expectation. Hence, if π^X is well-rounded (i.e., $R^2 \lesssim n$), then $\widetilde{O}(n^3 \operatorname{polylog} 1/\eta\varepsilon)$ queries suffice.*

This improves the prior best complexity of $\widetilde{O}(n^3R^2 \operatorname{polylog} \frac{1}{\varepsilon})$ for general logconcave sampling (in TV-distance) from scratch by [31, Collorary 1.2], and provides a much stronger $\mathcal{R}_\infty$-guarantee. Moreover, this complexity matches the best-known complexity for uniform sampling with the TV-guarantee by [9] and with the $\mathcal{R}_\infty$-guarantee by [23].

Result 3: Isotropic Rounding of Logconcave Distributions. The above results on the complexity of logconcave sampling still have dependence on the second moment R^2, so those are not fully polynomial in the problem parameters. We address this issue by initially running an algorithm for isotropic rounding which makes the covariance matrix of a given logconcave distribution near-isotropic (i.e., $\operatorname{Cov}\pi^X \approx I_n$). After this rounding, $R^2 = O(n)$, and thus the guarantee above turns into n^3 at the cost of a multiplicative factor of $\operatorname{polylog} R$ in the complexity. Isotropic rounding is also a useful tool for many other high-dimensional algorithms.

At a high level, just as in the warm-start generation, our algorithm follows a sequence $\{\mu_i\}$ of distributions while updating an affine map $F:\mathbb{R}^n \to \mathbb{R}^n$ along the way. An important subroutine is to convert a well-rounded distribution (i.e., with $\operatorname{tr}(\operatorname{Cov}) = O(n)$) to one that is near-isotropic (i.e., with $\operatorname{Cov} \approx I_n$). To this end, we generalize the approach taken in [12] for the uniform distribution — repeatedly [draw a few samples → compute a crude estimate of the covariance → identify skewed directions from the estimation and upscale them]. In doing so, we rely crucially on our sampler's mixing rate in terms of $\|\operatorname{Cov}\pi^X\|$ (instead of $\operatorname{tr}(\operatorname{Cov}\pi^X)$) in **Result 1** and its improved complexity for warm-start generation in **Result 2**.

Theorem 1.8. *For any logconcave distribution π^X specified by $\mathsf{Eval}_{x_0,R}(V)$, there exists a randomized algorithm with query complexity of $\widetilde{O}(n^{3.5} \operatorname{polylog} R)$ that finds an affine map F such that the pushforward of π^X via F is 1.01-isotropic with probability at least $1 - O(n^{-1/2})$.*

Using a boosting technique [34, Algorithm A], we can boost the success probability to $1-\eta$ by running the rounding algorithm $O(\log 1/\eta)$ more times. Moreover, our complexity improves upon the previous best complexity of $\widetilde{O}(n^4 \operatorname{polylog} R)$ for general logconcave distributions [31, Theorem 6.1]. Our rate also matches that of isotropic rounding for the uniform distribution [12, 13].

Putting the aforementioned components together, we can obtain an end-to-end guarantee for the query complexity of logconcave sampling given an evaluation oracle. To this end, we simply run the rounding algorithm (along with the boosting technique) and then Tilted Gaussian Cooling.

Corollary 1.9 (Polynomial complexity of logconcave sampling)**.** *For any $\eta,\varepsilon \in (0,1)$ and logconcave distribution π^X specified by $\mathsf{Eval}_{x_0,R}(V)$, there exists a randomized algorithm that with probability at least $1-\eta$ returns a sample Z such that $\mathcal{R}_\infty(\operatorname{law} Z \parallel \pi^X) \le \varepsilon$, using $\widetilde{O}(n^{3.5} \operatorname{polylog} \frac{R}{\eta} + n^3 \operatorname{polylog} 1/\eta\varepsilon)$ evaluation queries in expectation.*

Result 4: Integration of Logconcave Functions. Finally, we examine the complexity of integrating a logconcave function, a classical application of logconcave sampling. To this end, we extend an annealing approach in [9] to general logconcave integration. Similar to Tilted Gaussian Cooling, we follow a sequence $\{f_i\}_{i\in[m]}$ of logconcave functions, where f_1 is easy to integrate, $\mu_i \propto f_i$ is an $O(1)$-warm start for $\mu_{i+1} \propto f_{i+1}$, the variance of f_{i+1}/f_i with respect to μ_i is small enough (e.g., $\chi^2(\mu_{i+1} \parallel \mu_i) = O(m^{-1})$), and f_m is the target logconcave function.

Theorem 1.10. *For any $\varepsilon > 0$ and an integrable well-rounded logconcave function $f = e^{-V}:\mathbb{R}^n \to \mathbb{R}$ presented by $\mathsf{Eval}_{x_0,R}(V)$, there exists an algorithm that with probability at least $3/4$, returns a $(1+\varepsilon)$-multiplicative approximation to the integral of f using $\widetilde{O}(n^3/\varepsilon^2)$ queries. For an arbitrary logconcave f given by a well-defined function oracle, the complexity is bounded by $\widetilde{O}(n^{3.5} \operatorname{polylog} R + n^3/\varepsilon^2)$.*

This result improves the previous best complexity $\widetilde{O}(n^4/\varepsilon^2)$ of integrating logconcave functions [31, Theorem 1.3]. Again, our improved complexity matches that for the uniform distribution (i.e., volume computation) [9, 12, 13]. We note that due to the strong guarantees of the new sampler, we can streamline the analysis of errors and dependence among samples used to estimate f_{i+1}/f_i, simplifying and strengthening earlier analyses [9, 15, 33].

1.2 Techniques and Background

In proving our results — faster algorithms for the basic problems of sampling (with and without a warm start), rounding and integration — we do not take the most direct route. Instead we develop several techniques that appear to be interesting and we imagine will be useful in other contexts as well. These include the reversible heat-flow perspective for the design of polynomial-time algorithms providing a direct connection to isoperimetric constants, sampling guarantees in the strong notion of $\mathcal{R}_\infty$-distance (which has other motivations such as differential privacy), relevant geometry of logconcave functions, and a streamlined analysis of estimation errors of dependent samples.

1.2.1 Going beyond the Uniform Distribution. The uniform distribution over a convex body is a special case of a logconcave distribution. The best-known polynomial complexity for uniform sampling is achieved by addressing three subproblems: (1) sampling from a warm start, (2) generating a warm start and (3) performing isotropic rounding. Extending the best-known guarantees and algorithms for each of these problems to general logconcave densities is the primary challenged addressed in this paper. In our proofs, we use general properties of logconcave functions and avoid any structural assumptions on the target density.

The first challenge is to establish a logconcave sampler with a mixing rate matching that of the best uniform samplers, such as the Ball Walk [15] or In-and-Out [21]. These uniform samplers have a mixing rate of $n^2\|\mathrm{Cov}\,\pi^X\|$, while Hit-and-Run in [31], previously the best for general logconcave distributions, has a rate of $n^2\,\mathrm{tr}(\mathrm{Cov}\,\pi^X)$. The improved complexity bounds of these uniform samplers come from a better understanding of isoperimetric properties, such as the Cheeger and Poincaré constants for logconcave distributions, and this aspect is more direct for In-and-Out. Here we extend In-and-Out to general logconcave distributions. As we elaborate in §1.2.2, this requires implementing rejection sampling for the distributions of the form $\exp(-V(x)-\frac{1}{2h}\|x\|^2)$ (for a suitable h) using $\widetilde{O}(1)$ queries. However, with only logconcavity assumed for e^{-V}, determining a suitable proposal distribution and analyzing its query complexity present significant challenges.

Reduction to an Exponential Distribution. To extend methods from uniform to general logconcave distributions, we leverage a conceptual connection between logconcave sampling and convex optimization — sampling from $\exp(-V)$ is analogous to minimization of V. In optimization, we bound $V(x)$ with a new variable t, and add the convex constraint $V(x)\le t$. Inspired by this, we consider an exponential distribution with density $\pi(x,t)\propto\exp(-nt)\cdot\mathbb{1}[(x,t):V(x)\le nt]$ on $\mathbb{R}^{n+1}$, with X-marginal $\pi^X\propto\exp(-V)$. The scaling of n in the exponential function is a natural choice as we will see later.

This reduction offers several advantages. First, the potential V becomes linear, opening up a possible extension of previous analyses. Second, it points to a clear path to generalizing ideas/guarantees devised for uniform sampling. Specifically, the conditional law $\pi^{X|T=t}$ of X given $T=t$ is the uniform distribution over the convex level set $L_{nt}=\{V(x)\le nt\}$. Thus, π can be interpreted as an average of the uniform distribution over L_{nt} weighted by e^{-nt}, and this idea is crucial to our subsequent algorithms.

A similar exponential reduction was discussed conceptually in [18] and developed further in [19] for a sampling analogue of the interior-point method from optimization. Here, however, we apply this reduction in our general setup, without specific assumptions on the epigraph, and analyze structural properties of the reduced distribution. For example, we bound the mean and variance along the T-direction. This allows us to analyze how this reduction impacts key parameters of the original logconcave distribution, including the largest eigenvalue and trace of the covariance.

1.2.2 Logconcave Sampling from a Warm Start.

Sampling through Diffusion. The proximal sampler [26] is essentially a Gibbs sampler from two conditional distributions. For a target distribution $\pi^X\propto\exp(-V)$, it introduces a new variable Y and considers the augmented distribution $\pi^{X,Y}\propto\exp(-V(x)-\frac{1}{2h}\|x-y\|^2)$ with parameter $h>0$. One iteration involves two steps: (i) $y\sim\pi^{Y|X=x}=\mathcal{N}(x,hI_n)$ and (ii) $x\sim\pi^{X|Y=y}\propto\exp(-V(\cdot)-\frac{1}{2h}\|\cdot-y\|^2)$. While step (i) is straightforward, step (ii) requires a nontrivial sampling procedure.

The complexity of the Proximal Sampler involves a mixing analysis (to determine the number of iterations for a desired accuracy to the target) and the query complexity of implementing step (ii). For the first part, [7] demonstrated that one iteration corresponds to simulating the heat flow for (i) and then a *time reversal* of the heat flow for (ii). This leads to exponential decay of the $\mathcal{R}_q$-divergence, with the decay rate dependent on the isoperimetry of the target distribution (e.g, (PI) and (LSI)). For the second part, prior approaches typically use rejection sampling or a different logconcave sampler (e.g., MALA, ULMC) under additional assumptions such as the smoothness of V or access to first-order and proximal oracles for V.

Sampling without Smoothness. Previous studies of the Proximal Sampler focused on smooth unconstrained distributions, where the potential satisfies $\nabla^2 V\preceq\beta I_n$ for some $\beta<\infty$, leaving open the complexities of uniform or general logconcave sampling with hard constraints. [21] introduced In-and-Out, a version of the Proximal Sampler for uniform sampling over convex bodies. With $V(x)=(\mathbb{1}[x\in\mathcal{K}])^{-1}$ for a convex body $\mathcal{K}$, one iteration of In-and-Out draws $y\sim\mathcal{N}(x,hI_n)$ and then $x\sim\mathcal{N}(y,hI_n)|_{\mathcal{K}}$ (the Gaussian truncated to $\mathcal{K}$) using rejection sampling on the proposal $\mathcal{N}(y,hI_n)$. They introduce a *threshold* parameter on the number of rejection trials which ensures that the algorithm does not use too many queries.

This diffusion-based approach turns out to be stronger than the Ball Walk, the previous best uniform sampler with query complexity $\widetilde{O}(n^2\|\mathrm{Cov}\,\pi^X\|\,\mathrm{polylog}\,\frac{1}{\varepsilon})$ for obtaining an ε-close sample in TV-distance. In-and-Out provides $\mathcal{R}_q$-divergence guarantees with a matching rate of $n^2\|\mathrm{Cov}\,\pi^X\|$ for TV-distance, and its analysis is simpler than the Ball Walk. Since the latter's analysis in [15] goes through its biased version (called the Speedy Walk), it involves an understanding of an additional rejection step for making the biased distribution close to the uniform target, as well as the isoperimetric constant of the biased distribution. In contrast, In-and-Out achieves direct contraction towards the uniform target, with rate dependent on isoperimetric constants of the original target, not a biased one, and achieves stronger output guarantees. In-and-Out's approach

has also been extended to truncated Gaussian sampling [23], leading to a $\mathcal{R}_\infty$-guarantee that improves upon the TV-guarantee of the Ball Walk [9].

Proximal Sampler for General Logconcave Distributions. This prompts a natural question if we can extend In-and-Out beyond constant and quadratic potentials. Prior studies of the Proximal Sampler by [7, 21] allow for an immediate mixing analysis through the isoperimetry of logconcave distributions. However, we also need a query complexity bound for sampling from $\pi^{X|Y=y} \propto \exp(-V(x) - \frac{1}{2h}\|x-y\|^2)$. While one might consider using Hit-and-Run [31] for the second step, it requires *roundedness* of this distribution, and its current complexity is also unsatisfactory. Rejection sampling is another option, but without smoothness assumptions for general logconcave distributions, it is challenging to identify an appropriate proposal distribution and to bound the expected number of trials.

To handle an arbitrary convex potential V, we apply the exponential reduction, turning V into a convex constraint and work with the linear potential nt in one higher dimension. Extending In-and-Out to this exponential distribution, we propose the Proximal Sampler $\mathrm{PS}_{\exp}$ (Algorithm 1). Since $C_{\mathsf{PI}}(\pi) \lesssim (C_{\mathsf{PI}}(\pi^X)+1)\log n$, the mixing rate of $\mathrm{PS}_{\exp}$ for π is close to that of the Proximal Sampler for π^X. Implementing the second step is now simpler due to the linear potential, allowing us to extend the analysis of In-and-Out for uniform and truncated Gaussian distributions to exponential distributions. A new technical ingredient is bounding the rate of increase of $\int_{\mathcal{K}_\delta} e^{-nt} / \int_{\mathcal{K}} e^{-nt}$ as δ grows, where $\mathcal{K}_\delta = \{z \in \mathbb{R}^{n+1} : d(z,\mathcal{K}) \le \delta\}$. We show that this growth rate is bounded by $e^{\delta n}$.

1.2.3 Warm-start Generation. Prior work on warm-start generation [9, 15, 33] is based on using a sequence $\{\mu_i\}_{i\in[m]}$ of distributions, where μ_1 is easy to sample (e.g., uniform distribution over a unit ball), each μ_i is close to μ_{i+1} in probabilistic distance, and μ_m is the target. By moving along this sequence with a suitable sampler for intermediate annealing distributions, one can generate a warm start for the desired distribution. This approach is more efficient than trying to directly go from μ_1 to μ_m.

Warm-start Generation for Uniform Distributions. The state-of-the-art algorithm for uniform distributions over convex bodies is Gaussian Cooling [9]. They set $\mu_i(x) \propto \exp(-\|x\|^2/(2\sigma_i^2))$, where σ_i^2 increases from n^{-1} to R^2 according to a suitable schedule. This approach ensures that μ_i is $O(1)$-warm with respect to μ_{i+1}, and uses the Ball Walk to sample from truncated Gaussians with TV-guarantees. Using a coupling argument, they showed that with high probability, this scheme outputs an ε-close sample to the uniform distribution in TV-distance, using $\widetilde{O}(n^2(R^2\vee n)\,\mathrm{polylog}\,\frac{1}{\varepsilon})$ membership queries. [23] later improved this by replacing the Ball Walk with the Proximal Sampler, achieving $\mathcal{R}_\infty$-guarantees with the same complexity. As a result, Gaussian Cooling transfers $\mathcal{R}_\infty$-warmness across the sequence of distributions, and thus the complexity for uniform sampling from a convex body remains the same even for $\mathcal{R}_\infty$-divergence guarantees.

Going beyond Uniform Distributions. In [9], Cousins and Vempala raised the question of whether their annealing strategy can be extended to arbitrary logconcave distributions with complexity $\widetilde{O}(n^2(R^2\vee n))$. A natural choice for annealing distributions is $\mu(x) \propto \exp(-V(x) - \frac{1}{2\sigma^2}\|x\|^2)$, which would still provide $\mathcal{R}_\infty$-closeness of consecutive distributions and allow for accelerated updates to σ^2. However, prior samplers lack the necessary guarantees for these distributions, so we use the exponential reduction.

To generate a warm start for $\pi(x,t) \propto e^{-nt}|_{\mathcal{K}}$, it seems natural to consider an annealing distribution obtained by multiplying π by a Gaussian in "(x,t)" for a direct application of Gaussian Cooling in $\mathbb{R}^{n+1}$. However, due to different rates of changes in the quadratic term and linear term in t over an interval of length $O(1)$, these two terms do not properly cancel each other, which implies that the warmness of μ_{m-1} with respect to $\mu_m = \pi$ is no longer $O(1)$-bounded.

We address this by introducing $\mu_{\sigma^2,\rho}(x,t) \propto \exp(-\frac{1}{2\sigma^2}\|x\|^2 - \rho t)\cdot \mathbb{1}[(x,t)\in\bar{\mathcal{K}}]$ where $\sigma^2 \in (0,R^2]$, $\rho\in(0,n]$, and $\bar{\mathcal{K}} = \{V(x)\le nt\}\cap\{\|x\| = O(R), \|t\| = O(1)\}$. Essentially, this runs Gaussian Cooling along the x-direction with an exponential tilt in the t-direction, ensuring that consecutive distributions are $O(1)$-close in $\mathcal{R}_\infty$. However, we need an efficient sampler for $\mu_{\sigma^2,\rho}$ with $\mathcal{R}_\infty$-guarantees to maintain $\mathcal{R}_\infty$-warmness across the annealing scheme.

Sampling from Annealing Distributions. Given the form of the annealing distribution (the potential is a combination of linear and quadratic terms), we use the Proximal Sampler to develop $\mathrm{PS}_{\mathrm{ann}}$. The query complexity for rejection sampling in the second step now can be derived from our analysis of $\mathrm{PS}_{\exp}$ and the Proximal Sampler for truncated Gaussians [23]. For mixing with $\mathcal{R}_\infty$ guarantees, we apply a technique from [23]. We obtain a mixing rate of $O(h^{-1}C_{\mathsf{LSI}}(\mu_{\sigma^2,\rho}))$ for the Proximal Sampler based on (LSI). This results in only a *doubly logarithmic* dependence on the initial warmness, implying convergence from any *feasible* start with an overhead of $\mathrm{polylog}(n,R)$. This *uniform ergodicity* implies L^∞-norm contraction of density toward the target [10], leading to a $\mathcal{R}_\infty$-guarantee of $\mathrm{PS}_{\mathrm{ann}}$ without significant overhead.

We bound the LSI constant of the annealing distribution $\mu_{\sigma^2,\rho}$ by $\sigma^2\vee 1$ via the Bakry-Émery criterion and Holley–Stroock perturbation principle. For $\nu \propto \exp(-\frac{1}{2\sigma^2}\|x\|^2 - \frac{t^2}{2} - \rho t)|_{\bar{\mathcal{K}}}$, since the potential of ν is $\min(\sigma^{-2},1)$-strongly convex, its LSI without convex truncation is bounded by $\sigma^2\vee 1$ through Bakry-Émery, and convex truncation to $\bar{\mathcal{K}}$ only helps in satisfying the criterion [4]. Also, as $\sup t - \inf t = \Theta(1)$ over $\bar{\mathcal{K}}$, the ratio of $\mu_{\sigma^2,\rho}$ to ν is bounded below and above by $\Theta(1)$, so the perturbation principle ensures that $C_{\mathsf{LSI}}(\mu_{\sigma^2,\rho}) \lesssim C_{\mathsf{LSI}}(\nu) \le \sigma^2\vee 1$.

Gaussian Cooling with Exponential Tilt. With the query complexity of $\mathrm{PS}_{\mathrm{ann}}$ in mind, we design Tilted Gaussian Cooling for warm-start generation for $\pi(x,t)\propto\exp(-nt)|_{\mathcal{K}}$. We run rejection sampling with proposal $\mathcal{N}(0,n^{-1}I_n)\otimes\mathrm{Unif}(I_t)$ for some interval I_t of length $\Theta(1)$, and initial distribution $\exp(-\frac{n}{2}\|x\|^2)|_{\bar{\mathcal{K}}}$, which is $O(1)$-warm with respect to $\mu_1 = \exp(-\frac{n}{2}\|x\|^2 - t)|_{\bar{\mathcal{K}}}$. In Phase I, we update the two parameters according to $\sigma^2 \leftarrow \sigma^2(1+n^{-1})$ and $\rho\leftarrow\rho(1+n^{-1})$ while $\rho\le n$ and $\sigma^2\le 1$. Since Phase I involves $\widetilde{O}(n)$ inner phases with complexities of sampling from each annealing distributions being $\widetilde{O}(n^2(\sigma^2\vee 1)) = \widetilde{O}(n^2)$, the total complexity is $\widetilde{O}(n^3)$. In Phase II, we accelerate σ^2-updates via

$\sigma^2 \leftarrow \sigma^2(1+\sigma^2/R^2)$ as in Gaussian Cooling. With $\widetilde{O}(R^2/\sigma^2)$ inner phases (for doubling of σ^2) and sampling complexity $\widetilde{O}(n^2\sigma^2)$ per inner phase, this has total complexity $\widetilde{O}(n^2R^2)$. At termination, $\mathsf{PS}_{\exp}$ is run with $\mu_{R^2,n}$ as the initial distribution for target $\bar{\pi} \propto e^{-nt}|_{\bar{\mathcal{K}}}$, where these two are close in $\mathcal{R}_\infty$. Using the LSI of $\bar{\pi}$ and the boosting scheme again, we can achieve an ε-close sample to π (not $\bar{\pi}$) in the $\mathcal{R}_\infty$-divergence, using $\widetilde{O}(n^2(R^2 \vee n)\operatorname{polylog}\frac{1}{\varepsilon})$ queries in total.

1.2.4 Rounding. Rounding is the key to reducing the dependence on R from poly R to polylog R.

Isotropic Rounding for Uniform Distributions. The previous best rounding algorithm for uniform distributions, proposed by [12], gradually isotropizes a sequence $\{\mu_i\}$ of distributions. For a convex body $\mathcal{K}$, they set $\mu_i = \mathrm{Unif}(\mathcal{K} \cap B_{\delta^i}(0))$ for $\delta = 1+n^{-1/2}$, increasing i while $\delta^i \le R$. Their approach entails two important tasks: *Outer loop*: if $F(\mathcal{K} \cap B_r)$ is near-isotropic for an affine map F, then show that $F(\mathcal{K} \cap B_{\delta r})$ is well-rounded, and *Inner loop*: design an algorithm of n^3-complexity that isotropizes a given well-rounded uniform distribution. The first task was accomplished by [13] through Paouris' lemma (i.e., exponential tail decay) and a universal property that the diameter of an isotropic convex body is bounded by $n+1$.

The second task was addressed by repeating [draw samples → compute crude covariance estimation → upscale skewed directions of the covariance estimation]. They first run Gaussian Cooling to obtain a warm start for a uniform distribution μ from a convex body. Then, when the inner radius is r, the Ball Walk (or In-and-Out) is used to generate r^2 samples approximately distributed according to μ. These r^2 samples give a rough estimate of the covariance matrix $\overline{\Sigma}$ of μ such that $|\overline{\Sigma} - \Sigma| \precsim nI_n$, where $\Sigma = \operatorname{Cov}\mu$. Since the query complexity of these uniform samplers is $n^2\|\Sigma\|/r^2$, this procedure uses $n^2\|\Sigma\|$ queries in total. Then, it computes the eigenvalue/vectors of $\overline{\Sigma}$ and scales up (by a factor of 2) the subspace spanned by eigenvectors with eigenvalues less than n. One iteration of this process achieves two key properties: the largest eigenvalue of the covariance Σ increases by at most n additively while $r = \operatorname{inrad}\mu$ *almost* doubles. Since the well-roundedness ensures that $\|\Sigma\| = O(n)$ initially, the complexity of one iteration remains as $\widetilde{O}(n^3)$ throughout. Since there are $\sqrt{n}\log R$ outer iterations, the algorithm uses a total of $\widetilde{O}(n^{3.5}\operatorname{polylog} R)$ queries.

Extension to General Logconcave Distributions. Our rounding algorithm essentially follows this approach, with several technical refinements. First of all, we define a *ground set* of a general logconcave distribution, namely the level set $\mathsf{L}_g := \{x : V(x) - \min V \le 10n\}$. This ground set takes up most of measure due to the universal property in terms of the potential value. Focusing on the ground set is the first step toward a streamlined extension of the previous approach, so we consider the *grounded distribution* $\nu^X := \pi^X|_{\mathsf{L}_g}$. Then, for $\nu_r^X := \nu^X|_{B_r(0)}$, we isotropize a sequence of distributions, given by $\nu_1^X \to \nu_\delta^X \to \nu_{\delta^2}^X \to \cdots \to \nu_D^X \to \nu^X \to \pi^X$ for $D = \Theta(R)$.

To analyze the outer loop, we show that for an affine map F^X between $\mathbb{R}^n$, if $F^X_\# \nu_r^X$ is near-isotropic, then $(4F^X)_\# \nu_{\delta r}^X$ is well-rounded. Unfortunately, a varying density of $F^X_\# \nu_{\delta r}^X$ poses a daunting challenge in extending the previous proof in [13] to general logconcave distributions. Nonetheless, we can resolve this issue by working again with the exponential reduction. This extension asks for a universal property that an isotropic grounded distribution has diameter of order $O(n)$, similar to isotropic uniform distributions. We show this in the full paper. Transferring roundedness in the steps $\nu_D^X \to \nu^X$ and $\nu^X \to \pi^X$ is relatively straightforward by combining the change of measures and the reverse Hölder inequality for logconcave distributions.

For the inner loop, we can still apply the algorithm from [12] (or its streamlined version in [22]) with only minor changes to constants. First, with the logconcave sampler $\mathsf{PS}_{\exp}$ whose mixing rate depends on $\|\operatorname{Cov}\pi^X\|$ rather than $\operatorname{tr}(\operatorname{Cov}\pi^X)$, the complexity analysis of the inner loop extends naturally to general logconcave distributions. Next, we note that the proofs for controlling $\|\Sigma_i\|$ and $\operatorname{tr}\Sigma_i$ are identical to those for uniform distributions. The proof for the doubling of the inner radius at each iteration is nearly the same as in the uniform case, since the existence of a large ball due to isotropy is also a *universal* property of logconcave distributions.

1.2.5 Integration. For integration, we use the stronger guarantees of our logconcave sampler, along with a version of the Tilted Gaussian Cooling scheme to obtain a cubic algorithm for *well-rounded* logconcave functions. For general logconcave functions, we use the rounding algorithm as a pre-processing step, then apply the integration algorithm to the near-isotropic distribution obtained after rounding.

Volume Computation through Annealing. Similar to sampling, prior volume algorithms also follow a sequence $\{f_i\}_{i\in[m]}$ of logconcave functions, moving across distributions $\mu_i \propto f_i$ using a logconcave sampler. The annealing scheme is designed in a way that f_1 is easy to integrate, μ_i is an $O(1)$-warm start for μ_{i+1}, the variance of the estimator $E_i = f_{i+1}/f_i$ with respect to μ_i is bounded by $m^{-1}\mathbb{E}_{\mu_i}[E_i^2]$ (i.e., $\chi^2(\mu_{i+1} \,\|\, \mu_i) = O(m^{-1})$), and $f_{m+1} = f$ is the target logconcave function. Since $\mathbb{E}_{\mu_i}E_i = \int f_{i+1}/\int f_i$, accurate estimations of all E_i guarantee that the product $\int f_1 \times E_1 \cdots E_m$ is a good estimator of $\int f_1 \cdot \mathbb{E}E_1 \cdots \mathbb{E}E_m = \int f$.

The best-known algorithm, Gaussian Cooling [9], uses unnormalized Gaussian densities $f_i(x) = \exp(-\|x\|^2/2\sigma_i^2) \cdot \mathbb{1}_{\mathcal{K}}(x)$ for a convex body $\mathcal{K}$, along with an update schedule for σ_i^2. As described earlier in §1.2.3, its uses the Ball Walk to sample from $\mu_i \propto f_i$ with TV-guarantees.

Extension to General Logconcave Functions. Our integration algorithm follows this approach, but once again in the lifted space. With $\mathsf{PS}_{\mathrm{ann}}$ used for sampling, we follow a modified version of Tilted Gaussian Cooling for ease of analysis, particularly for variance control. We use $f_i(x,t) = \exp(-\|x\|^2/2\sigma_i^2 - \rho_i t)|_{\mathcal{K}}$ as the intermediate annealing functions, and ensure that $\mathcal{R}_\infty(\mu_i \,\|\, \mu_{i+1}) = O(1)$ for efficient sampling from μ_{i+1}, and $\mathcal{R}_2(\mu_{i+1} \,\|\, \mu_i) = \exp(O(m^{-1}))$ for $\mathrm{var}_{\mu_i}(E_i/\mathbb{E}_{\mu_i}E_i) = O(m^{-1})$. Since $O(1)$-warmness can be shown as in Tilted Gaussian Cooling, we elaborate on technical tools for variance control along with design of the algorithm. In Phase I, we go from $\exp(-n\|x\|^2/2)|_{\mathcal{K}}$ to $\exp(-\|x\|^2/2)|_{\mathcal{K}}$ with the update $\sigma^2 \leftarrow \sigma^2(1+1/n)$, where variance control is achieved by the logconcavity of $a \mapsto a^n \int h^a$ for a logconcave function h [14]. In Phase II, we move from $\exp(-\|x\|^2/2 - t)|_{\mathcal{K}}$ to $\exp(-\|x\|^2/2 - nt)|_{\mathcal{K}}$ with the update $\rho \leftarrow \rho(1+1/n)$, where variance control follows from the previous lemma and another lemma in [9]. Lastly in Phase

III, we move from $\exp(-\|x\|^2/2 - nt)|_{\mathcal{K}}$ to $\exp(-nt)|_{\mathcal{K}}$ with the update $\sigma^2 \leftarrow \sigma^2(1+\sigma^2/n)$, and we use the lemma in [9] again for variance control.

Streamlined Statistical Analysis. If all samples used are *independent,* and the estimators have moderate variance, namely $\mathrm{var}_{\mu_i}(E_i/\mathbb{E}E_i) = O(m^{-1})$, then $\mathrm{var}(E_1\cdots E_m) = (\prod(1+m^{-1}) - 1)\,(\mathbb{E}E_i)^2 \approx O(1)\,(\mathbb{E}[E_1\cdots E_m])^2$, which implies concentration of the estimator around $\mathbb{E}[E_1\cdots E_m] = \mathbb{E}E_1\cdots\mathbb{E}E_m$ through Chebyshev's inequality. However, samples given by (say) the Ball Walk in Gaussian Cooling are *approximately* distributed according to μ_i. In fact, samples drawn from μ_i and μ_{i+1} are *dependent*, since a sample from μ_i is used as a warm-start for μ_{i+1}. This means that the estimator $E_1\cdots E_m$ is a *biased estimator* of $\mathbb{E}E_1\cdots\mathbb{E}E_m (\neq \mathbb{E}[E_1\cdots E_m])$. These two issues complicate statistical analysis in prior work on how close the estimator is to the integral of f, and require additional technical tools to address them. For the first issue, previous work used a coupling argument based on TV-distance (referred to as "divine intervention") to account for the effects of approximate distributions (rather than exactly μ_i). For the second, they used the notion of α-mixing [35] (referred to as "μ-independence" therein) to bound the bias of the product estimator.

We simplify this statistical analysis substantially by using stronger guarantees of our sampler $\mathsf{PS}_{\mathsf{ann}}$. For the first issue, when $\bar{\mu}_i$ denotes the actual law of a sample X_i satisfying $\mathcal{R}_\infty(\bar{\mu}_i \,\|\, \mu_i) \le \varepsilon$, we can notice that the probability of a bad event (any event) with respect to $\bar{\mu}_i$ (instead of μ_i) only increases by at most a multiplicative factor of $1+\varepsilon$. Since the mixing time of $\mathsf{PS}_{\mathsf{ann}}$ has a polylogarithmic dependence on $1/\varepsilon$, we can set ε polynomially small (i.e., $\varepsilon \leftarrow \varepsilon/\mathrm{poly}(n,R)$), so we can enforce $\mathcal{R}_\infty(\otimes_i\bar{\mu}_i\|\otimes_i\mu_i) \le \varepsilon$ without a huge overhead in the query complexity. For the second issue, we use notion of β-mixing (or the coefficient of absolute regularity) [17], which is stronger than α-mixing. This quantity basically measures the discrepancy between a joint distribution and the product of marginal distributions by $\|\mathrm{law}(X_i, X_{i+1}) - \mathrm{law}\,X_i \otimes \mathrm{law}\,X_{i+1}\|_{\mathsf{TV}} = \|\mathrm{law}(X_i, X_{i+1}) - \bar{\mu}_i \otimes \bar{\mu}_{i+1}\|_{\mathsf{TV}} \le \beta$. Since the mixing of $\mathsf{PS}_{\mathsf{ann}}$ via (LSI) ensures mixing from *any start*, we can easily bound β by $O(\varepsilon)$. Thus, when analyzing the probability of a bad event, we can replace $\otimes\bar{\mu}_i$ with $\mathrm{law}(X_1,\dots,X_m)$ at the additive cost of $O(m\varepsilon)$ in probability (we can replace $\varepsilon \leftarrow m\varepsilon$ once again).

1.3 Definitions and Notation

Let $\mathcal{P}(\mathbb{R}^n)$ be the family of probability measures (distributions) on $\mathbb{R}^n$ that are absolutely continuous with respect to the Lebesgue measure. We use the same symbol for a distribution and density. For a set S and its indicator function $\mathbb{1}_S(x) = [x \in S]$, we use $\mu|_S$ to denote a distribution μ truncated to S (i.e., $\mu|_S \propto \mu\cdot\mathbb{1}_S$). For a measurable map $F : \mathbb{R}^n \to \mathbb{R}^n$ and $\mu \in \mathcal{P}(\mathbb{R}^n)$, the *pushforward* measure $F_\#\mu$ is defined as $F_\#\mu(A) = (\mu\circ F^{-1})(A)$ for a measurable set A. For $a, b \in \mathbb{R}$, we use $a\vee b$ and $a\wedge b$ to indicate their maximum and minimum, respectively. We use $B_r^n(x_0)$ to denote the n-dimensional ball of radius $r>0$ centered at $x_0 \in \mathbb{R}^n$, dropping the superscript n if there is no confusion. We use $a = \widetilde{O}(b)$ to denote $a = O(b\,\mathrm{polylog}\,b)$. For two symmetric matrices A, B, we use $|A| \preceq B$ to denote $-B \preceq A \preceq B$. Unless specified otherwise, for a vector v and a PSD matrix M, $\|v\|$ and $\|\Sigma\|$ refer to the ℓ_2-norm of v and the operator norm of Σ, respectively. We use $\overline{\mathbb{R}}$ to denote the extended real number system $\mathbb{R}\cup\{\pm\infty\}$.

We recall notions of common probability divergences/distances between distributions.

Definition 1.11. For $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$, the *$f$-divergence* of μ towards ν with $\mu \ll \nu$ is defined as, for a convex function $f : \mathbb{R}_+ \to \mathbb{R}$ with $f(1) = 0$ and $f'(\infty) = \infty$,

$$D_f(\mu \,\|\, \nu) := \int f\Big(\frac{\mathrm{d}\mu}{\mathrm{d}\nu}\Big)\,\mathrm{d}\nu\,.$$

For $q \in (1,\infty)$, the KL-*divergence and χ^q-divergence* correspond to $f(x) = x\log x$ and $x^q - 1$, respectively. The *q-Rényi divergence* is defined as

$$\mathcal{R}_q(\mu \,\|\, \nu) := \tfrac{1}{q-1}\log\big(\chi^q(\mu \,\|\, \nu) + 1\big)\,.$$

The *Rényi-infinity divergence* is defined as

$$\mathcal{R}_\infty(\mu \,\|\, \nu) := \log \operatorname{ess\,sup}_\mu \frac{\mathrm{d}\mu}{\mathrm{d}\nu}\,.$$

A distribution μ is said to be *M-warm with respect to a distribution ν* if $\frac{\mu(S)}{\nu(S)} \le M$ for any measurable subset S (i.e., μ is $\exp(\mathcal{R}_\infty(\mu\,\|\,\nu))$-warm with respect to ν). The *total variation* (TV) distance for $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ is defined by

$$\|\mu-\nu\|_{\mathsf{TV}} := \frac{1}{2}\int |\mu(x)-\nu(x)|\,\mathrm{d}x = \sup_{S\in\mathcal{F}} |\mu(S)-\nu(S)|\,,$$

where $\mathcal{F}$ is the collection of all measurable subsets of $\mathbb{R}^n$.

We recall $\mathsf{KL} = \lim_{q\downarrow 1}\mathcal{R}_q \le \mathcal{R}_q \le \mathcal{R}_{q'} \le \mathcal{R}_\infty$ for $1 \le q \le q'$ and $2\,\|\cdot\|_{\mathsf{TV}}^2 \le \mathsf{KL} \le \mathcal{R}_2 = \log(\chi^2+1) \le \chi^2$. We refer readers to [36] for basic properties of the Rényi-divergence (e.g., continuity/monotonicity in q).

ACKNOWLEDGMENTS

We are grateful to Andre Wibisono for helpful discussions and to Sinho Chewi for pointing us to the Holley-Stroock principle and the reference [38]. This work was supported in part by NSF Award CCF-2106444 and a Simons Investigator Award.

References

[1] Christophe Andrieu, Anthony Lee, Sam Power, and Andi Q. Wang. 2024. Explicit convergence bounds for Metropolis Markov chains: isoperimetry, spectral gaps and profiles. *Ann. Appl. Probab.* 34, 4 (2024), 4022–4071.
[2] David Applegate and Ravi Kannan. 1991. Sampling and integration of near log-concave functions. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing (STOC '91)*. ACM Press, 156–163.
[3] Filippo Ascolani, Hugo Lavenant, and Giacomo Zanella. 2024. Entropy contraction of the Gibbs sampler under log-concavity. *arXiv preprint arXiv:2410.00858* (2024).
[4] Dominique Bakry, Ivan Gentil, and Michel Ledoux. 2014. *Analysis and geometry of Markov diffusion operators*. Vol. 348. Springer, Cham.
[5] Peter Buser. 1982. A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup. (4)* 15, 2 (1982), 213–230.
[6] Jeff Cheeger. 1970. A lower bound for the smallest eigenvalue of the Laplacian. In *Problems in analysis (Sympos. in honor of Salomon Bochner, Princeton Univ., Princeton, N.J., 1969)*. Princeton Univ. Press, Princeton, NJ, 195–199.
[7] Yongxin Chen, Sinho Chewi, Adil Salim, and Andre Wibisono. 2022. Improved analysis for a proximal algorithm for sampling. In *Conference on Learning Theory (COLT)*. PMLR, 2984–3014.
[8] Sinho Chewi, Murat A Erdogdu, Mufan Li, Ruoqi Shen, and Shunshi Zhang. 2022. Analysis of Langevin Monte Carlo from Poincaré to log-Sobolev. In *Conference on Learning Theory (COLT)*. PMLR, 1–2.
[9] Ben Cousins and Santosh Vempala. 2018. Gaussian cooling and $O^*(n^3)$ algorithms for volume and Gaussian volume. *SIAM J. Comput.* 47, 3 (2018), 1237–1273.

[10] Pierre Del Moral, Michel Ledoux, and Laurent Miclo. 2003. On contraction properties of Markov kernels. *Probab. Theory Related Fields* 126, 3 (2003), 395–420.
[11] Martin Dyer, Alan Frieze, and Ravi Kannan. 1991. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. Assoc. Comput. Mach.* 38, 1 (1991), 1–17.
[12] He Jia, Aditi Laddha, Yin Tat Lee, and Santosh Vempala. 2021. Reducing isotropy and volume to KLS: an $O^*(n^3\psi^2)$ volume algorithm. In *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing (STOC '21)*. ACM, New York, 961–974.
[13] He Jia, Aditi Laddha, Yin Tat Lee, and Santosh S Vempala. 2024. Reducing isotropy and volume to KLS: Faster rounding and volume algorithms. *arXiv preprint arXiv:2008.02146* (2024).
[14] Adam Tauman Kalai and Santosh Vempala. 2006. Simulated annealing for convex optimization. *Math. Oper. Res.* 31, 2 (2006), 253–266.
[15] Ravi Kannan, László Lovász, and Miklós Simonovits. 1997. Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Structures Algorithms* 11, 1 (1997), 1–50.
[16] Bo'az Klartag. 2023. Logarithmic bounds for isoperimetry and slices of convex sets. *Ars Inven. Anal.* (2023), Paper No. 4, 17.
[17] Andrei Nikolaevich Kolmogorov and Yu A. Rozanov. 1960. On a strong mixing condition for stationary Gaussian processes. *Teor. Verojatnost. i Primenen.* 5 (1960), 222–227.
[18] Yunbum Kook, Yin Tat Lee, Ruoqi Shen, and Santosh S Vempala. 2022. Sampling with Riemannian Hamiltonian Monte Carlo in a constrained space. In *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 35. 31684–31696.
[19] Yunbum Kook and Santosh S Vempala. 2024. Gaussian Cooling and Dikin Walks: The Interior-Point Method for Logconcave Sampling. In *Conference on Learning Theory (COLT)*. PMLR, 3137–3240.
[20] Yunbum Kook and Santosh S. Vempala. 2024. Sampling and integration of logconcave functions by algorithmic diffusion. *arXiv preprint arXiv:2411.13462* (2024).
[21] Yunbum Kook, Santosh S. Vempala, and Matthew S. Zhang. 2024. In-and-Out: Algorithmic diffusion for sampling convex bodies. In *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 37. Curran Associates, Inc., 108354–108388.
[22] Yunbum Kook and Matthew S Zhang. 2024. Covariance estimation using Markov chain Monte Carlo. *arXiv preprint arXiv:2410.17147* (2024).
[23] Yunbum Kook and Matthew S Zhang. 2025. Rényi-infinity constrained sampling with d^3 membership queries. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*. SIAM, 5278–5306.
[24] Michel Ledoux. 1994. A simple analytic proof of an inequality by P. Buser. *Proc. Amer. Math. Soc.* 121, 3 (1994), 951–959.
[25] Michel Ledoux. 2004. Spectral gap, logarithmic Sobolev constant, and geometric bounds. In *Surveys in differential geometry. Vol. IX*. Surv. Differ. Geom., Vol. 9. Int. Press, Somerville, MA, 219–240.
[26] Yin Tat Lee, Ruoqi Shen, and Kevin Tian. 2021. Structured logconcave sampling with a restricted Gaussian oracle. In *Conference on Learning Theory (COLT)*. PMLR, 2993–3050.
[27] Yin Tat Lee and Santosh S. Vempala. 2024. Eldan's stochastic localization and the KLS conjecture: isoperimetry, concentration and mixing. *Ann. of Math. (2)* 199, 3 (2024), 1043–1092.
[28] László Lovász. 1991. *How to compute the volume?* DIMACS, Center for Discrete Mathematics and Theoretical Computer Science.
[29] László Lovász and Miklós Simonovits. 1990. The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. In *1990 IEEE 31st Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE Comput. Soc. Press, Los Alamitos, CA, 346–354.
[30] László Lovász and Miklós Simonovits. 1993. Random walks in a convex body and an improved volume algorithm. *Random Structures Algorithms* 4, 4 (1993), 359–412.
[31] László Lovász and Santosh Vempala. 2006. Fast Algorithms for Logconcave Functions: Sampling, Rounding, Integration and Optimization. In *2006 IEEE 47th Annual Symposium on Foundations of Computer Science (FOCS)*. 57–68.
[32] László Lovász and Santosh Vempala. 2006. Hit-and-Run from a corner. *SIAM J. Comput.* 35, 4 (Jan. 2006), 985–1005.
[33] László Lovász and Santosh Vempala. 2006. Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. *J. Comput. System Sci.* 72, 2 (2006), 392–417.
[34] László Lovász and Santosh Vempala. 2007. The geometry of logconcave functions and sampling algorithms. *Random Structures Algorithms* 30, 3 (2007), 307–358.
[35] Murray Rosenblatt. 1956. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. U.S.A.* 42 (1956), 43–47.
[36] Tim van Erven and Peter Harremoës. 2014. Rényi divergence and Kullback-Leibler divergence. *IEEE Trans. Inform. Theory* 60, 7 (2014), 3797–3820.
[37] Santosh S. Vempala and Andre Wibisono. 2023. Rapid convergence of the unadjusted Langevin algorithm: isoperimetry suffices. In *Geometric Aspects of Functional Analysis*. Lecture Notes in Math., Vol. 2327. Springer, Cham, 381–438.
[38] Feng-Yu Wang. 2013. *Analysis for diffusion processes on Riemannian manifolds*. Vol. 18. World Scientific.

Received 2024-11-04; accepted 2025-02-01