Hit-and-run mixing via localization schemes

Yuansi Chen[1] and Ronen Eldan[2]

[1]Duke University
[2]Microsoft Research, Redmond

Abstract

We analyze the hit-and-run algorithm for sampling uniformly from an isotropic convex body K in n dimensions. We show that the algorithm mixes in time $\tilde{O}(n^2/\psi_n^2)$, where ψ_n is the smallest isoperimetric constant for any isotropic logconcave distribution, also known as the Kannan-Lovasz-Simonovits (KLS) constant [KLS95]. Our bound improves upon previous bounds of the form $\tilde{O}(n^2R^2/r^2)$, which depend on the ratio R/r of the radii of the circumscribed and inscribed balls of K, gaining a factor of n in the case of isotropic convex bodies. Consequently, our result gives a mixing time estimate for the hit-and-run which matches the state-of-the-art bounds for the ball walk. Our main proof technique is based on an annealing of localization schemes introduced in Chen and Eldan [CE22], which allows us to reduce the problem to the analysis of the mixing time on truncated Gaussian distributions.

1 Introduction

Sampling from a high dimensional distribution is a fundamental computational problem in many fields, such as Bayesian statistics, machine learning, statistical physics, and others involving stochastic models. A particularly important class of high dimensional distributions consists of uniform distributions over convex bodies. For example, the problem of sampling uniformly from a convex body is closely related to that of efficiently computing its volume, which is a fundamental problem in computer science and has been extensively studied in the last three decades (see [DFK91, LS90, AK91, LS93], the survey [Vem05] and the thesis [Cou17]). Besides the volume computation, uniform sampling from a convex body can be seen as a special case of sampling from truncated Gaussian distributions, which arise naturally in Bayesian statistical models involving probit regression and censored data [AC93, HH06].

The *hit-and-run* algorithm is a widely-used Markov-chain-based sampling method. It was introduced by Smith in 1984 [Smi84] and is closely related to the popular Gibbs sampler [Tur71]. In the case of uniform sampling from a convex body, the hit-and-run algorithm works iteratively as follows. At each step, it starts from a point x inside the convex body, chooses a uniformly distributed random direction, and then samples a point uniformly from the line segment formed by intersection of the convex body with the line segment passing through x in the randomly chosen direction. This process is repeated until the chain is well-mixed.

The hit-and-run algorithm has been shown to mix rapidly for any log-concave distribution [LV06a]. In particular, the works [Lov99, LV06b] show that the hit-and-run algorithm mixes in $\tilde{O}(n^2R^2/r^2)$ from a warm start or any interior point, where R and r are the radii of the circumscribed and inscribed balls of K. In the case of *isotropic* convex bodies, it was an open question whether the mixing time of the hit-and-run has to depend on the ratio R/r, which is typically of the order $\sqrt{n}$ (see [LV18b]). By isotropic, we mean that the uniform distribution over the convex body has mean 0 and covariance $\mathbb{I}_n$. Consequently, it was unknown whether the hit-and-run algorithm mixes in $\tilde{O}(n^2)$ steps for any isotropic convex body. The main contribution of this paper is to provide a positive answer to this question:

Theorem 1 (main theorem, informal). *Let μ be the uniform distribution on an n-dimensional isotropic convex set K. Let ν be a measure obtained by running t steps of the hit-and-run chain starting from any*

measure whose density with respect to μ is at most M. Then the total variation distance between μ and ν is at most ϵ under the condition

$$t \geq n^2 \left(\frac{M \log(n)}{\epsilon} \right)^{O(1)}.$$

A formal statement of this result can be found in Theorem 2 at the end of this section. Our result makes two main contributions to the existing literature on sampling from convex sets, as described next.

First, our result effectively matches the known mixing bound for the hit-and-run walk with the state-of-the-art bound for the *ball walk*, another well-studied sampling algorithm. Given a convex set K from which we want to produce samples, the ball walk is the Markov chain whose step is defined as follows: Given a starting point x, it chooses a point y uniformly in a ball of fixed radius around x and moves to y if y is in K; otherwise, it rejects the move and stays at x. It was shown in [KLS97] that the ball walk mixes in $\tilde{O}(n^2 R^2/r^2)$ from a warm start, where R and r are the radii of the circumscribed and inscribed balls of K. In addition to the result which depends on the ratio R/r, the results in [KLS97] also imply that the ball walk mixes in $\tilde{O}(n^2/\psi_n^2)$ steps for an isotropic convex body, where ψ_n is the Kannan-Lovasz-Simonovits (KLS) constant [KLS95]. Using the best known bound for ψ_n [Che21, KL22, JLV22] which is of order $\log^{-5} n$, the mixing time of the ball walk becomes $\tilde{O}(n^2)$. Thus, compared to the ball walk, our result gives a matching bound in terms of the dimension n dependency for sampling an isotropic convex body using the hit-and-run. To the best of our knowledge, the ball walk is currently the Markov chain with the best known mixing rate for sampling a general isotropic convex body.

Second, our result is the first application of the *annealing with localization schemes* technique, which was put forth in [CE22], towards sampling from continuous distributions. The main proof strategy is to use the stochastic localization process [Eld13] in order to reduce the original mixing time analysis to that of sampling from a *truncated Gaussian distribution*, which is known to be well-behaved in many cases. While the high-level of the proof follows this technique, the adaptation of the technique to the hit-and-run chain seems to require substantial additional work, resulting in a framework which we hope would become relevant for other chains as well.

Furthermore, in regard to the comparison with the ball walk, we would like to highlight that unlike the ball walk which depends on a parameter indicating the jump size (which corresponds to the radius of the ball), the hit-and-run chain is more canonical in the sense that it does not depend on a "scale" parameter. Given a membership oracle for an isotropic convex body, it is often not clear what the optimal choice of jump size should be, with no extra information about the convex body one needs to assume the worst case (taking a jump size of $n^{-1/2}$). The hit-and-run chain, by construction, chooses the (in a sense) optimal jump size for each point. Therefore, in light of our bound, while for the worst-case example of an isotropic convex body the two algorithms have comparable performance, there are cases of convex bodies for which the hit-and-run algorithm is strictly faster than the ball-walk (this is essentially true whenever the inscribed radius is much bigger than order 1). In other words our bound effectively establishes that the hit-and-run walk is always at least as good as the ball walk for sampling for isotropic convex bodies, but in some cases it's actually strictly better.

Remark. It is known that the hit-and-run also mixes rapidly from a *cold start*, namely it can be started from any single interior point of the convex body [LV06b]. Specifically, [LV06b] shows that the mixing time in this case is $\tilde{O}\left(n^2 \frac{R^2}{r^2} \text{polylog}(M)\right)$, hence the dependence on the warmness parameter M is poly-logarithmic rather than polynomial. This leads to the open question of whether one can obtain a $\tilde{O}(n^2 \text{polylog}(M))$ mixing bound for the hit-and-run chain in the case of isotropic convex bodies.

1.1 Related work

Sampling uniformly from a convex body is a well-studied problem in the literature. The first polynomial-time algorithm for this problem was proposed by Dyer, Frieze and Kannan [DFK91]. Many subsequent algorithms and improved mixing times have been developed [LS93, KLS97]. The best known mixing

time for sampling a general convex body, whose circumscribed and inscribed balls have radii R and r, respectively, from a warm start is $\tilde{O}(n^2 R^2 / r^2)$, achieved by the ball walk in [KLS97]. The results in [KLS97] also imply that the ball walk mixes in $\tilde{O}(n^2/\psi_n^2)$ for an isotropic convex body, where ψ_n is the Kannan-Lovasz-Simonovits (KLS) constant [KLS95]. Sampling from a convex body is closely related to the problem of efficiently computing its volume. Faster mixing often leads to faster volume computation algorithms. For related literature on volume computation, we refer the readers to [KLS97, LV06c, CV18, JLLV21] and the references therein.

Since a convex polytope is a special case of a convex body, the problem of sampling uniformly from a polytope is covered by algorithms which sample from general convex bodies. However, the additional structure of a polytope also allows for new algorithms with provably better mixing times [KN09, LV17, LV18a, CDWY18, MV19, LLV20].

Specifically for the hit-and-run, [Lov99] shows that it mixes rapidly from a constant-warm start. That is, in $\tilde{O}(n^2 R^2 / r^2)$ steps, the total variation distance between the current distribution and the stationary distribution is at most a small constant. Here R and r are the radii of the circumscribed and inscribed balls of the convex body, respectively. This bound has the same order of magnitude as for the ball walk, so hit-and-run is no worse when we assume a bound on R/r.

While the ball walk is known to mix slowly if started at a corner of the convex body, the hit-and-run is known to mix rapidly from any single interior point. [LV06b] shows that the hit-and-run mixes in $\tilde{O}(n^2 R^2 / r^2 \text{polylog}(M))$ from any M-warm start. So even if M is of order n to a fixed degree, the mixing time remains in the same order. This line of work has been extended in [LV06a] for sampling general logconcave distributions with smoothness assumptions.

In terms of proof techniques, our proof uses the stochastic localization process [Eld13] to transform the original uniform distribution to a simpler truncated Gaussian distribution. The idea of using localization processes towards proving mixing bounds for Markov chains is summarized in the framework introduced in [CE22]. Localization schemes have also been applied to the analysis of high dimensional distributions that arise in functional analysis, convex and discrete geometry, combinatorics and mathematical physics (see the survey [Eld22]). In particular, the use of the stochastic localization process has led to the near-resolution of the Kannan, Lovász and Simonovits conjecture, Bourgain's hyperplane conjecture and the thin-shell conjecture in convex geometry [Che21, KL22, JLV22].

1.2 Formal statement of the problem and results

Next, we make the required definitions towards the precise statement of our main result. We assume that $K \subset \mathbb{R}^n$ is a convex body, which is a compact convex set with nonempty interior.

Target distribution. The distribution from which we want to sample is the uniform measure on the convex body $K \subset \mathbb{R}^n$,

$$\mu(x) \propto \mathbf{1}_K(x), \forall x \in \mathbb{R}^n.$$

Isotropic position. We say a measure ν on $\mathbb{R}^n$ is *isotropic* if

$$\mathbb{E}_{X \sim \nu}[X] = 0 \text{ and } \text{Var}_{X \sim \nu}[X] = \mathbb{I}_n.$$

A convex body is called isotropic if the corresponding uniform measure is isotropic.

Hit-and-run Markov chain for a convex body. The *hit-and-run* chain is defined by the following transition step: given the current state $u \in K$, we generate unit vector $\theta \in \mathbb{R}^n$ which is sampled from the uniform measure on the unit sphere and consider the line $\ell := \{u + t\theta; \ t \in \mathbb{R}\}$. The next point is then chosen uniformly from the segment $\ell \cap K$.

Hit-and-run-transition kernel for a general target density. Next, we give a more general definition of the hit-and-run chain. Given a density ν with respect to the Lebesgue measure on $\mathbb{R}^n$, we denote by $P_{u\to\cdot}(\nu)$ the hit-and-run one-step transition kernel with starting point $u \in \mathbb{R}^n$ with respect to the underlying (target) density ν, which is defined as follows: For any measurable set $A \subseteq \mathbb{R}^n$, we have

$$P_{u\to A}(\nu) := \frac{2}{n\pi_n} \cdot \int_A \frac{\nu(x)dx}{\nu(\ell_{ux})\,|u-x|^{n-1}}, \tag{1}$$

where π_n is the volume of the unit ball $\pi_n = \mathrm{vol}(\mathbb{B}^n) = \frac{\pi^{n/2}}{\mathrm{Gamma}(n/2+1)}$, with Gamma as the gamma function and $\nu(\ell_{ux})$ is the integral of ν along the line ℓ_{ux} through u and x. Specifically, for ℓ_{ux} the line through u and x, we define

$$\nu(\ell_{ux}) := \int_{\ell_{ux}} \nu(v)\mathcal{H}_1(dv)$$

where $\mathcal{H}_1$ is the one-dimensional Hausdorff measure. It is straightforward to check that when taking ν to be the uniform measure on K, this definition identifies with the previous one. Additionally, it is not hard to see that the above chain is reversible, and its stationary distribution is ν [LV06b].

Lazy chain. Given a Markov chain with transition kernel P. We define its lazy variant $P^{\text{after-lazy}}$, which stays in the same state with probability at least $\frac{1}{2}$, as

$$P_{x\to S}^{\text{after-lazy}} = \frac{1}{2}\delta_{x\to S} + \frac{1}{2}P_{x\to S}.$$

Here $\delta_{x\to\cdot}$ is the Dirac distribution at x. Since the lazy variant only slows down the convergence rate by a constant factor, we study lazy Markov chains in this paper for its convenience in theoretical analysis.

Next we introduce a few notions in order to quantify the mixing time of the hit-and-run algorithm.

Total-variance distance. We denote the total variation (TV) distance between two probability distributions $\mathcal{P}_1, \mathcal{P}_2$ by

$$\mathrm{d_{TV}}(\mathcal{P}_1, \mathcal{P}_2) = \sup_{A\in\mathfrak{B}(\mathbb{R}^n)} |\mathcal{P}_1(A) - \mathcal{P}_2(A)|,$$

where $\mathfrak{B}(\mathbb{R}^n)$ is the Borel sigma-algebra on $\mathbb{R}^n$. If $\mathcal{P}_1$ and $\mathcal{P}_2$ admit densities p_1 and p_2 respectively, we may write

$$\mathrm{d_{TV}}(\mathcal{P}_1, \mathcal{P}_2) = \frac{1}{2}\int |p_1(x) - p_2(x)|\,dx.$$

Warm start. We say an initial distribution μ_{init} is M-*warm* if it satisfies

$$\sup_{S\in\mathfrak{B}(\mathbb{R}^n)} \frac{\mu_{\text{init}}(S)}{\mu(S)} \le M.$$

If M is a constant that does not depend on n, then we say μ_{init} is *constant M-warm*.

Mixing time. For an error tolerance $\epsilon \in (0,1)$, the total variance distance ϵ-mixing time of the Markov chain P with initial distribution μ_{init} and target distribution μ is defined as

$$\mathrm{t_{mix}}(\epsilon, \mu_{\text{init}}, \mu) := \inf\left\{ k \in \mathbb{N} \mid \mathrm{d_{TV}}\left(\mathcal{T}_P^k(\mu_{\text{init}}), \mu\right) \le \epsilon \right\}.$$

With the above definitions in hand, we state our main theorem. We are interested in obtaining an upper bound for the mixing time of the hit-and-run algorithm for sampling from an isotropic density $\mu \propto \mathbf{1}_K$ in terms of the ϵ-mixing time. Our main theorem reads:

Theorem 2. *Let μ be the uniform distribution on an n-dimensional isotropic convex body $K \subset \mathbb{R}^n$. There exist universal constants $C, c > 0$, such that for any M-warm initial distribution μ_{init} and any error tolerance $\epsilon \in (0, 1)$ such that $n \geq c \log \frac{M}{\epsilon}$, the ϵ-mixing time of the lazy hit-and-run is upper bounded as follows*

$$\mathsf{t}_{\mathrm{mix}}\left(\epsilon, \mu_{\mathrm{init}}, \mu\right) \leq C \frac{n^2}{\psi_n^2} \left(\frac{M}{\epsilon}\right)^{11} \log^5 \frac{M}{\epsilon}.$$

The proof, together with an outline of the proof strategy, are provided in Section 3.

2 Preliminaries

In this section, we introduce notation, background and preliminary results needed for our proof.

2.1 Logconcavity and concentration

Logconcave density.   We say a density ν is *logconcave* if it satisfies

$$\nu(x)^\tau \nu(y)^{1-\tau} \leq \nu(\tau x + (1 - \tau)y), \text{for } x, y \in \mathbb{R}^n, \tau \in [0, 1].$$

For example, $\mu \propto \mathbf{1}_K$ with K being a convex set is logconcave.

Cheeger's isoperimetric constant.   We define the *Cheeger's isoperimetric constant* of a measure ν

$$\psi_\nu := \inf_{A \subseteq \mathbb{R}^n} \left\{ \frac{\int_{\partial A} \nu}{\min\left\{\nu(A), 1 - \nu(A)\right\}} \right\}.$$

And

$$\psi_n := \inf_{\nu \text{ isotropic logconcave on } \mathbb{R}^n} \psi_\nu.$$

It is known that $\psi_n \geq \log^{-5}(n)$ [KL22]. A closely related quantity if $\kappa_n > 0$ defined as follows

$$\kappa_n^2 := \sup_{\nu \text{ isotropic logconcave}} \sup_{\theta \in \mathbb{S}^n} \left\{ \left|\mathbb{E}_{X \sim \nu} \langle X, \theta \rangle \left(X \otimes X\right)\right|^2 \right\},$$

where the first supremum is taken over all isotropic logconcave measures on $\mathbb{R}^n$ and $\mathbb{S}^n$ is the unit sphere in $\mathbb{R}^n$. It is known in Eldan [Eld13] that there exists a universal constant $C > 0$ such that $\frac{1}{\psi_n^2} \leq C \log n \cdot \kappa_n^2$.

2.2 Definitions related to geometric convexity

Let K be a convex body (compact, convex, full-dimensional convex set) in $\mathbb{R}^n$. Denote by $\mathbb{B}^n(x, \tau)$ with ball with center x and radius τ in $\mathbb{R}^n$. Let vol be the n-dimensional Lebesgue measure. Define

$$\lambda(u, t) := \frac{\mathrm{vol}(K \cap \mathbb{B}^n(u, t))}{\mathrm{vol}(\mathbb{B}^n(0, t))}, \tag{2}$$

the fraction of a ball of radius t centered around u that intersects K. Let K_r be the set of points $x \in K$ with large $\lambda(x, 2r)$, that is,

$$K_r := \left\{ x \in K \mid \lambda\left(x, 2r\right) \geq \frac{63}{64} \right\}. \tag{3}$$

As shown in [Lov99], the set K_r is convex thanks to the Brunn-Minkowski Theorem.

2.3 Definitions regarding Markov chains

We are interested in the problem of sampling from a target measure ν on $\mathbb{R}^n$. Given a Markov chain with transition kernel $P : \mathbb{R}^n \times \mathfrak{B}(\mathbb{R}^n) \to \mathbb{R}_{\geq 0}$ where $\mathfrak{B}(\mathbb{R}^n)$ denotes the Borel σ-algebra on $\mathbb{R}^n$, the k-step transition kernel P^k is defined recursively by

$$P^k_{x \to dy} = \int_{z \in \mathbb{R}^n} P^{k-1}_{x \to dz} P_{z \to dy}.$$

Associated transition operator. Let $\mathcal{T}_P$ denote the transition operator associated to the Markov chain. It is defined as

$$\mathcal{T}_P(\nu)(S) := \int_{y \in \mathbb{R}^n} d\nu(y) P(y, S), \quad \forall S \in \mathfrak{B}(\mathbb{R}^n).$$

When ν is the distribution of the current state, $\mathcal{T}_P(\nu)$ is the distribution of the next state. And $\mathcal{T}_P^n(\nu) := \mathcal{T}_{P^n}(\nu)$ is the distribution of the state after n steps.

Dirichlet form. Let $\mathcal{L}_2(\nu)$ be the space of square integrable functions under the measure ν. The *Dirichlet form* $\mathcal{E}_P : \mathcal{L}_2(\nu) \times \mathcal{L}_2(\nu) \to \mathbb{R}_{\geq 0}$ associated with the transition kernel P is given by

$$\mathcal{E}_P(f, g) := \frac{1}{2} \int \int \left(f(x) - f(y) \right) \left(g(x) - g(y) \right) P_{x \to dy} d\nu(x).$$

Truncated conductance. For $s \in (0, 1)$, we define the s-conductance Φ_s of the Markov chain P with its stationary measure ν as follows

$$\Phi_s(P) := \inf_{S : s < \nu(S) < 1-s} \frac{\int_S P_{x \to S^c} d\nu(x)}{\min\{\nu(S), \nu(S^c)\} - s}. \tag{4}$$

When compared to conductance (the case $s = 0$), s-conductance allows us to ignore small parts of the distribution where the conductance is difficult to bound. Remark that using the Dirichlet form notation, we can write

$$\Phi_s(P) = \inf_{S : s < \nu(S) < 1-s} \frac{\mathcal{E}(\mathbf{1}_S, \mathbf{1}_S)}{\min\{\nu(S), \nu(S^c)\} - s}.$$

The following lemma by Lovász and Simonovits [LS93] connects the s-conductance with the mixing time of a Markov chain.

Lemma 3 (Corollary 1.5 in [LS93])**.** *Consider a reversible lazy Markov chain with transition kernel P and stationary distribution μ. Let μ_{init} be an M-warm initial distribution. Let $0 < s < \frac{1}{2}$. Then*

$$\mathrm{d}_{\mathrm{TV}}\left(\mathcal{T}_P^N(\mu_{\text{init}}), \mu \right) \leq Ms + M \left(1 - \frac{\Phi_s^2}{2} \right)^n.$$

2.4 Background on the stochastic localization process

For a probability density ν on $\mathbb{R}^n$, define

$$\mathbf{b}(\nu) := \int x \nu(x) dx,$$

its center of mass.

Given a density μ on $\mathbb{R}^n$, we define the stochastic localization (SL) process ([Eld13]) with a positive semi-definite control matrix C_t, by

$$\mu_t(x) := \frac{1}{Z(t, c_t)} \exp\left(c_t^\top x - \frac{1}{2}\langle B_t x, x\rangle \right) \mu(x), \tag{5}$$

where c_t and B_t satisfy the following stochastic differential equations:

$$dc_t = C_t dW_t + C_t^2 \mathbf{b}(\mu_t) dt$$
$$dB_t = C_t^2 dt.$$

Here W_t is the standard Brownian motion and C_t is any process adapted to W_t which takes values in the space of $n \times n$ matrices. The existence and uniqueness of the solutions of the SDE is shown via the standard existence and uniqueness results on SDEs (see e.g. Lemma 3 in [Che21]).

It is known that μ_t satisfies the following SDE for $x \in \mathbb{R}^n$

$$d\mu_t(x) = (x - \mathbf{b}(\mu_t))^\top C_t dW_t \mu_t(x). \tag{6}$$

Define also

$$A_t := \int (x - \mathbf{b}(\mu_t))(x - \mathbf{b}(\mu_t))^\top \mu_t(x) dx, \tag{7}$$

the covariance matrix of μ_t.

Note that running the SL process with a starting measure $\mu \propto \mathbf{1}_K$ results in a density at time t which is a random density which has an explicit form of a truncated Gaussian on a convex set.

Truncated Gaussian on a convex set. For $m > 0$ and $\beta \in \mathbb{R}^n$, define

$$\nu_{\beta,m}(x) := \frac{e^{-\frac{m}{2}|x-\beta|^2} \mathbf{1}_K}{\int_{\mathbb{R}^n} e^{-\frac{m}{2}|x-\beta|^2} \mathbf{1}_K dx}.$$

It is the Gaussian with mean $\beta \in \mathbb{R}^n$ and variance $\frac{1}{m}$ supported on the convex set K.

In light of Eq. (5), we have that for all $t \geq 0$, under the choice $C_s = \mathbb{I}_n$ for all $s \in [0, t]$, the measure μ_t obtained by the SL process has the form

$$\mu_t = \nu_{c_t/t, t}. \tag{8}$$

The following is a special case of a classical result by Brascamp and Lieb [BL02]:

Theorem 4. *One has* $\mathrm{Cov}(\nu_{\beta,m}) \preceq \frac{1}{m}\mathbb{I}_n$.

If follows that, for every $t > 0$, assuming the choice $C_s = \mathbb{I}_n$ for $s \in [0, t)$, one has almost surely

$$A_t \preceq \frac{1}{t}\mathbb{I}_n. \tag{9}$$

2.5   Other notation

We use $\gamma : \mathbb{R} \to \mathbb{R}_+$ to denote the standard Gaussian density function, and $\Gamma : \mathbb{R} \to [0, 1]$ to denote the standard Gaussian cumulative density function. We use $\mathcal{N}(b, \sigma^2)$ to denote the Gaussian measure with mean b and variance σ^2. For p and $\tilde{p}$ two measures, $p * \tilde{p}$ denotes the convolution of the two.

We use big-O notation $O(\cdot)$ to denote asymptotic upper bounds which ignore all constants. For example, we write $g_1(n) = O(g_2(n))$ if there exists a universal constant $c > 0$ such that $g_1(n) \leq cg_2(n)$ when n is larger than a universal constant. We use $\tilde{O}(\cdot)$ to denote asymptotic upper bounds which ignore both constants and poly-logarithmic factors on the parameters involved.

3 Proof of the main theorem

We prove Theorem 2, by bounding the s-conductance, via Lemma 3. Roughly speaking, for constant M-warm initial distribution, taking

$$s = \frac{\epsilon}{2M}, n \geq \frac{2}{\Phi_s^2} \log \frac{2M}{\epsilon}$$

results in a mixing time of $\frac{2}{\Phi_s^2} \log \frac{2M}{\epsilon}$. As a consequence, the main focus of the proof is in lower bounding the s-conductance.

Unlike [Lov99], we do not bound the s-conductance directly, as that requires an argument that depends in rather intricate ways on the geometry of the convex set K. Instead, we employ an annealing of localization schemes introduced in [CE22] which attempts to reduce the analysis to a simpler case in which the measure is localized, in the sense that the uniform measure on K is multiplied by a Gaussian density with small variance.

This is done by considering the stochastic localization (SL) process defined in subsection 2.4 to the measure μ. Given $\mu \sim \mathbf{1}_K$, we consider the process $(\mu_t)_{t \geq 0}$ defined by equation (5) with the choice $C_t = \mathbb{I}_n$ up to time T. We fix the choice of the time $T = n$.

The annealing technique boils down to the following two main steps:

1. Show that for a fixed set E whose measure is bounded away from 0 and 1, the quantity $\mu_t(E)$ is also bounded away from those values with non-negligible probability. This behavior is referred to in [CE22] as *approximate conservation of variance*. Under this condition, the conductance for the transition kernel at time 0 can be lower-bounded in terms of the conductance of the transition kernel at time T, so that one only needs to give a lower bound on the latter quantity.

2. Bound the conductance of the hit-and-run chain which corresponds to the measure μ_T which, according to Eq. (8), is a Gaussian with variance $\frac{1}{T}\mathbb{I}_n$ restricted to the set K. This is an easier task than the analysis of the original conductance, since this measure is typically localized well-inside the convex body K, so that a step of hit-and-run is hardly affected by its boundary.

The first of the two steps highlighted above is captured by the following lemma.

Lemma 5. *Let K be an isotropic convex body in $\mathbb{R}^n$ and let μ_T be defined as above. Let $\zeta > 0$ and let K_r be a convex subset of K with $\mu(K_r) \geq 1 - \zeta/100$. Then for any $E \subset K_r$ whose measure $\mu(E) = \xi \in (0, 1/2]$ satisfies $\zeta \leq c\xi^2/\sqrt{\log(10^4/\xi)}$, we have that*

$$\mathbb{P}\left(\mu_T(E)(\mu_T(K_r) - \mu_T(E)) \geq \frac{1}{16} \cdot \mu(E)(\mu(K_r) - \mu(E))\right) \geq \frac{c\xi^{3/2}}{\log \frac{1}{\xi}\sqrt{n \log(n)\kappa_n^2}},$$

for a universal constant $c > 0$.

The following lemma lower bounds the volume of K_r used in Lemma 5. See [Lov99] for a proof.

Lemma 6 (Lemma 2 in [Lov99]). *Suppose K contains a unit ball. Let $\mu \propto \mathbf{1}_K$. Then*

$$\mu(K_r) \geq 1 - 2\sqrt{n}r.$$

In order to reduce the conductance of the original Markov chain to that of $P_{\cdot \to \cdot}(\mu_T)$, we also need the following lemma. Its proof follows from [CE22, Proposition 48] and from the fact that the hit-and-run chain is the Markov chain associated to the subspace-localization scheme described in [CE22, Section 2.1].

Lemma 7. *Consider the process $(\mu_t)_t$ defined above. Fix a function $f \in \mathcal{L}_2(\mu)$. Let $P_t = P_{\cdot \to \cdot}(\mu_t)$. Define*

$$D_t := \mathcal{E}_{P_t}(f, f).$$

Then $(D_t)_{t \geq 0}$ is a super-martingale.

The next lemma, which corresponds to the second step described above, gives the conductance for the hit-and-run on the "transformed" density μ_T.

Lemma 8. *There exists a universal constant $c > 0$ such that the following holds true. Let $\nu_{\beta,n}$ be a probability measure defined as a truncated Gaussian on a convex set, given by the formula*

$$\nu_{\beta,n}(x) \propto e^{-\frac{n}{2}|x-\beta|^2} \mathbf{1}_K.$$

Define $\Upsilon := \left\{ u \in K \mid |u - \beta| \in \left[\frac{1}{\sqrt{2}}, \sqrt{2} \right] \right\}$ and $\delta := 1 - \nu_{\beta,n}(\Upsilon)$. Suppose $S_1 \cup S_2$ is a partition of K and let $0 < r \leq \frac{1}{16\sqrt{n}}$. Then we have

$$\int_{S_1} P_{u \to S_2}(\nu_{\beta,n}) d\nu_{\beta,n}(x) \geq \frac{r^2\sqrt{n}}{c} \left[\nu_{\beta,n}(S_1 \cap K_r) \cdot \nu_{\beta,n}(S_2 \cap K_r) - 8 \left(1 + \frac{32}{r} \right) \delta \right].$$

Recall from Eq. (8) that μ_t is uniquely determined by the vector c_t, and is given by the formula $\mu_t = \nu_{c_t/t,t}$. The next lemma shows that at time $T = n$, the set Υ defined in Lemma 8 has large measure with high probability.

Lemma 9. *Let $(\mu_t)_t$ be the process defined above and let $(c_t)_t$ be the corresponding process which appears in Eq. (5), then at time $T = n$,*

(i) *the random vector c_n/n has the law $\mu * \mathcal{N}(0, \frac{1}{n}\mathbb{I}_n)$.*

(ii) *there exists an event $\mathfrak{E} \subseteq \mathbb{R}^n$ with measure at least $1 - 2e^{-\frac{n}{32}}$ under the law of c_n/n, such that for $v \in \mathfrak{E}$, we have*

$$\mathbb{P}_{\mathbf{x} \sim \nu_{v,n}} \left(\frac{\sqrt{2}}{2} < |\mathbf{x} - v| < \sqrt{2} \right) \geq 1 - e^{-\frac{n}{32}}.$$

Proof of Theorem 2. Fix the following parameters

$$\begin{aligned}
s &= \frac{\epsilon}{2M}, \\
\zeta &= s^2/\sqrt{\log(10^4/s)}/10^8, \\
r &= \frac{\zeta}{200\sqrt{n}}, \\
T &= n.
\end{aligned} \tag{10}$$

Let $(\mu_t)_t$ be the stochastic localization process applied to the measure μ. As highlighted above, we provide a lower bound for the s-conductance of μ in terms of that of μ_T, and then derive a lower bound for the s-conductance of μ_T.

According to Theorem 4.1 in [KLS95], for K isotropic in $\mathbb{R}^n$, there exists a point $x \in K$ such that $\mathbb{B}^n(x, 1) \subseteq K$. Together with Lemma 6, for the above choice of r, we obtain

$$\mu(K_r) \geq 1 - \frac{\zeta}{100}.$$

Let $S_1 \cup S_2 = K$ be a partition of K, with $s \leq \mu(S_1) \leq \frac{1}{2}$. Let $P_t = P_{\cdot \to \cdot}(\mu_t)$ be the hit-and-run transition kernel with the target density μ_t. To lower bound the s-conductance, we need to lower bound

$$\mathcal{E}_{P_0}(\mathbf{1}_{S_1}, \mathbf{1}_{S_1}) = \int_{S_1} P_{x \to S_2}(\mu) d\mu(x).$$

Define

$$\mathfrak{E} = \left\{ v \in \mathbb{R}^n \mid \mathbb{P}_{\mathbf{x} \sim \nu_{v,n}} \left(\frac{\sqrt{2}}{2} < |\mathbf{x} - v| < \sqrt{2} \right) \geq 1 - e^{-\frac{n}{32}} \right\}.$$

According to Lemma 9 we have

$$\mathbb{P}\left(\frac{c_n}{n} \in \mathfrak{E}\right) \geq 1 - 2e^{-\frac{n}{32}}. \tag{11}$$

Applying Lemma 8 with r as above, $\beta = c_n/n$ and $\delta = e^{-n/32}$, we obtain

$$\int_{S_1} P_{u \to S_2}(\mu_n) d\mu_n(x) \geq \frac{r^2\sqrt{n}}{c'''}\left(\mu_n(S_1 \cap K_r) \cdot \mu_n(S_2 \cap K_r) - 8\left(1 + \frac{32}{r}\right)e^{-n/32}\right)\mathbf{1}_{\frac{c_n}{n}\in\mathfrak{E}}$$

$$\geq \left(\frac{c\zeta^2}{\sqrt{n}}\mu_n(S_1 \cap K_r) \cdot \mu_n(S_2 \cap K_r) - c'\sqrt{n}e^{-n/32}\right)\mathbf{1}_{\frac{c_n}{n}\in\mathfrak{E}}. \tag{12}$$

We have

$$\int_{S_1} P_{x \to S_2}(\mu) d\mu(x)$$

$$\overset{(i)}{\geq} \mathbb{E}\left[\int_{S_1} P_{x \to S_2}(\mu_n) d\mu_n(x)\right]$$

$$\overset{(12)}{\geq} \mathbb{E}\left[\left(\frac{c\zeta^2}{\sqrt{n}}\mu_n(S_1 \cap K_r) \cdot \mu_n(S_2 \cap K_r) - c'\sqrt{n}e^{-n/32}\right)\mathbf{1}_{\frac{c_n}{n}\in\mathfrak{E}}\right]$$

$$\overset{(ii)}{\geq} \frac{c\zeta^2}{\sqrt{n}}\left(\frac{c''s^{3/2}}{\log\frac{1}{s}\sqrt{n\log(n)}\kappa_n^2} - 2e^{-n/32}\right) \cdot \mu(S_1 \cap K_r) \cdot (\mu(K_r) - \mu(S_1 \cap K_r)) - c'\sqrt{n}e^{-n/32}$$

$$\overset{(iv)}{\geq} C \cdot \frac{s^{5.5}}{n\log^2(\frac{1}{s})\kappa_n\sqrt{\log(n)}}\left[\min\{\mu(S_1), \mu(S_1^c)\} - s/2\right],$$

where $c, c'c'', c''', C$ are all universal constants. Here (i) follows from Lemma 7, which claims that $(\mathcal{E}_{P_t}(\mathbf{1}_{S_1}, \mathbf{1}_{S_1}))_{t\geq 0}$ is a super-martingale. (ii) follows from Lemma 5 and Eq. (11). (iv) follows because there exists a constant $c > 1$ such that for $n \geq c\log(\frac{M}{\epsilon})$, meaning that $\sqrt{n\log(n)}\kappa_n e^{-n/32} \ll s^2$. The above calculation establishes a lower bound on the s-conductance

$$\Phi_s \geq C\frac{s^{5.5}}{n\log^2(\frac{1}{s})\kappa_n\sqrt{\log(n)}}.$$

According to [KL22], $\frac{1}{\psi_n^2} \leq C\log(n)\kappa_n^2$. We conclude with Lemma 3. $\qquad\square$

Overview of the rest of the proof In the rest of the proof, in Section 4 we prove Lemma 5, which shows the approximate conservation of variance of any indicator function. In Section 5 we prove Lemma 8 which shows the conductance of the hit-and-run on the transformed density, and Lemma 9 which shows that with high probability most of the mass of the transformed density is concentrated in a shell.

4   Approximate conservation of variance

The goal of this section is to prove Lemma 5. To this end, we fix a set $E \subset \mathbb{R}^n$; our aim is to show that we non-negligible probability we have both that $\mu_T(E)$ is bounded away from 0 and 1 and that $\mu_T(K_r)$ is close to 1.

We have not been able to show this directly with respect to the SL process with the choice $C_t = \mathbb{I}_n$. Instead, we consider a different choice of driving matrix C_t and stopping time (described below) which on one hand makes the analysis more tractable and on the other hand has the resulting distribution of the random measure being the same as that of μ_T.

4.1 Construction of the three-stage process

The driving matrix C_t is chosen so that the process has three different stages, as follows.

Given $\mu \sim \mathbf{1}_K$, $E \subset \mathbb{R}^n$ such that $\mu(E) = \xi \in (0, 1/2]$, we run three stages of SL to obtain $(\hat{\mu}_t)_{t \geq 0}$ as follows

- **Stage 1:** Starting from μ, run SL with $C_t = I_n$ from time 0 to time $T_1 := \frac{\xi}{\kappa_n^2 \log(n)}$ to obtain $(\hat{\mu}_t)_{t \in [0, T_1]}$.

- **Stage 2:** Starting from $\hat{\mu}_{T_1}$, we choose a driving matrix C_t which satisfies the following. There is a stopping time T_2 so that:

    1. One has $\hat{\mu}_{T_1}(E) = \hat{\mu}_{T_2}(E)$ almost surely.
    2. The matrix $n\mathbb{I}_n - \int_0^{T_2} C_t^2 dt$ is positive definite and its rank is almost surely at most 1.

- **Stage 3:** Starting from $\hat{\mu}_{T_2}$, and defining $n\mathbb{I}_n - \int_0^{T_2} C_t^2 dt = \lambda_1 \theta \theta^\top$ where $|\theta| = 1$, run SL with $C_t = \theta \theta^\top$ for a time period $n - \lambda_1$.

Note that the driving matrix in Stage 2 is defined only implicitly. The exact choice of driving matrix which satisfies the two conditions of this stage is of no consequence for the rest of the proof, rather it is only important to establish to existence of such a choice. Roughly speaking these two conditions can be obtained by choosing $C_t = \mathrm{Proj}_{H_t \cap v_t^\perp}$, where $v_t = \int_E (x - \mathbf{b}(\hat{\mu}_t)) d\hat{\mu}_t(x)$ and H_t is the image of the matrix $n\mathbb{I}_n - \int_0^t C_s^2 ds$. We refer the reader to [EKZ21, Lemma 2] for the exact construction.

Observe that it follows from the definition of the process that, at the end of Stage 3, one has almost surely

$$\int_0^{T_3} C_t^2 dt = n\mathbb{I}_n. \tag{13}$$

Since $\hat{\mu}_t$ is a martingale and T_3 is a stopping time, applying the optional stopping theorem yields

$$\mathbb{E}[\hat{\mu}_{T_3}(x)] = \mu(x), \quad \forall x \in \mathbb{R}^n. \tag{14}$$

According to formula 8, there exists a random variable $\hat{\mathbf{y}}_{T_3}$ on $\mathbb{R}^n$ such that $\hat{\mu}_{T_3}$ takes the form

$$\hat{\mu}_{T_3}(x) = \nu_{\hat{\mathbf{y}}_{T_3}, n}(x) = \frac{1}{Z_{T_3}(\hat{\mathbf{y}}_{T_3})} \exp\left(-\frac{n}{2} |x - \hat{\mathbf{y}}_{T_3}|^2\right) \mu(x), \quad \forall x \in \mathbb{R}^n.$$

where $Z_{T_3}(\hat{\mathbf{y}}_{T_3})$ is the normalizing constant. Letting $p_{\hat{\mathbf{y}}_{T_3}}$ be the distribution of $\hat{\mathbf{y}}_{T_3}$, Eq. (14) implies that

$$\int_{\mathbb{R}^n} \nu_{y,n}(x) p_{\hat{\mathbf{y}}_{T_3}}(dy) = \mu(x), \quad \forall x \in \mathbb{R}^n.$$

On the other hand, applying the same argument for μ_T which was obtained via the SL process with driving matrix $C_t = \mathbb{I}_n$, there exists a random variable $\mathbf{y}_n$ such that, almost surely

$$\mu_T(x) = \frac{1}{Z(\mathbf{y}_n)} \exp\left(-\frac{n}{2} |x - \mathbf{y}_n|^2\right) \mu(x) = \nu_{\mathbf{y}_n, n}(x), \quad \forall x \in \mathbb{R}^n.$$

Denoting the law of $\mathbf{y}_n$ by $p_\mathbf{y}$, the martingale property yields

$$\int_{\mathbb{R}^n} \nu_{y,n}(x) p_\mathbf{y}(dy) = \mu(x), \quad \forall x \in \mathbb{R}^n.$$

We conclude that

$$\int_{\mathbb{R}^n} \nu_{y,n}(x) p_{\hat{\mathbf{y}}_{T_3}}(dy) = \int_{\mathbb{R}^n} \nu_{y,n}(x) p_\mathbf{y}(dy) = \mu(x), \quad \forall x \in \mathbb{R}^n. \tag{15}$$

The following lemma shows that $p_{\hat{\mathbf{y}}_{T_3}} = p_\mathbf{y}$, which implies that μ_T and $\hat{\mu}_{T_3}$ have the same distribution. Its proof is provided in Subsection 4.3.

Lemma 10. *Given $\mu(x) \sim \mathbf{1}_K$ the uniform distribution on a bounded convex set $K \subset \mathbb{R}^n$. If there exists a density p on $\mathbb{R}^n$, such that*

$$\mu(x) = \int \nu_{y,n}(x)p(y)dy, \forall x \in \mathbb{R}^n,$$

then p is uniquely defined almost everywhere.

Since the statement of Lemma 5 only depends on the random measure μ_T and is oblivious of the path leading to it, we can prove this lemma via the analysis of the process $(\hat{\mu}_t)_{t \in [0,T_3]}$. The rest of the proof therefore boils down to the next lemma, proven in Subsection 4.2.

Lemma 11. *Suppose that $\mu = \mathbf{1}_K$ is isotropic and logconcave and that $\mu(K_r) \geq 1 - \zeta/100$. Let $E \subset K_r$ satisfy $\mu(E) = \xi \in (0, 1/2]$ with $\zeta \leq \xi^2/\sqrt{\log(10^4/\xi)}/10^8$. Let $(\hat{\mu}_t)_t$ be the 3-stage process defined above. Then,*

$$\mathbb{P}\left(\hat{\mu}_{T_3}(E)(\hat{\mu}_{T_3}(K_r) - \hat{\mu}_{T_3}(E)) \geq \frac{1}{16} \cdot \mu(E)(\mu(K_r) - \mu(E))\right) \geq \frac{c\xi^{3/2}}{\log(\frac{1}{\xi})\sqrt{n\kappa_n^2 \log(n)}},$$

for a universal constant $c > 0$.

Proof of Lemma 5. Lemma 5 directly follows from (15) combined with the two lemmas above. $\square$

4.2 Approximate conservation of variance for the process $\hat{\mu}_t$

To prove Lemma 11, we proceed with three lemmas which deal with each stage of the stochastic process one by one. Consider the following events,

$$W_1 := \left\{\hat{\mu}_{T_1}(E) \in \left[\xi - \frac{\xi}{4}, \xi + \frac{\xi}{4}\right]\right\} \cap \left\{\hat{\mu}_{T_1}(K_r) \geq 1 - \frac{\zeta}{20}\right\},$$

$$W_2 := \left\{\hat{\mu}_{T_2}(E) \in \left[\xi - \frac{\xi}{4}, \xi + \frac{\xi}{4}\right]\right\} \cap \left\{\hat{\mu}_{T_2}(K_r) \geq 1 - \frac{\zeta}{10}\right\}, \text{ and}$$

$$W_3 := \left\{\hat{\mu}_{T_3}(E) \in \left[\frac{\xi}{2}, \frac{3}{4}\right]\right\} \cap \left\{\hat{\mu}_{T_3}(K_r) \geq \frac{7}{8}\right\}.$$

Lemma 12. *For $\mu(E) = \xi \in (0, 1/2]$ and $\mu(K_r) \geq 1 - \zeta/100$, there exists a universal constant $C > 0$ such that at the end of Stage 1 (meaning, for $T_1 = \frac{\xi}{C\kappa_n^2 \log n}$), we have*

$$\mathbb{P}(W_1) \geq 0.6.$$

Lemma 13. *We have*

$$\mathbb{P}(W_2 | W_1) \geq 0.5.$$

Lemma 14. *Under the assumption $\zeta \leq \xi^2/\sqrt{\log(10^4/\xi)}/10^8$, we have*

$$\mathbb{P}(W_3 | W_2) \geq \frac{c}{\log(\frac{1}{\xi})}\sqrt{\frac{T_1}{n}},$$

for a universal constant $c > 0$.

4.2.1 Analysis of Stage 1

The proof of Lemma 12 relies on the analysis developed in recent years around the KLS conjecture (see e.g. [Che21, KL22]). We apply the following upper bound on the operator norm of the covariance matrix A_t, proven by Klartag and Lehec:

Lemma 15 (Lemma 5.2 in [KL22]). *For every $T \leq (C\kappa_n^2 \log n)^{-1}$ we have*

$$\mathbb{P}(\|A_t\|_2 \geq 2 \, for \, 0 \leq t \leq T) \leq \exp\left(-\frac{1}{CT}\right),$$

where C is a universal constant.

While Lemma 5.2 in [KL22] only shows the result for a fixed $t \leq T$, it is not hard to see that, using Doob's inequality, the same proof can be generalized to the case of all $t \in [0, T]$ with a small modification on the constant C.

Equipped with Lemma 15, Lemma 12 then follows from a simple stochastic calculus.

Proof of Lemma 12. Let $g_t = \hat{\mu}_t(E)$. We have

$$dg_t = \int_E (x - \mathbf{b}(\hat{\mu}_t))^\top dW_t \hat{\mu}_t(x) dx.$$

Its quadratic variation is

$$d[g]_t = \left| \int_E (x - \mathbf{b}(\hat{\mu}_t)) \hat{\mu}_t(x) dx \right|^2 dt$$
$$\leq \|A_t\|_2 \, dt,$$

where A_t is the covariance matrix of $\mathbf{b}(\hat{\mu}_t)$. We have

$$\mathbb{P}\left(\hat{\mu}_t(E) \in \left[\xi - \frac{\xi}{4}, \xi + \frac{\xi}{4}\right]\right) = \mathbb{P}\left(\tilde{W}_{[g]_t} \in \left[-\frac{\xi}{4}, \frac{\xi}{4}\right]\right)$$
$$\geq 0.9 - \mathbb{P}\left(\int_0^t \|A_\tau\|_2 \, d\tau > \frac{\xi}{64}\right)$$

Applying Lemma 15 on the operator norm of A_t with $t \leq \frac{\xi}{C\kappa_n^2 \log n}$, it follows that

$$\mathbb{P}\left(\hat{\mu}_{T_1}(E) \in \left[\xi - \frac{\xi}{4}, \xi + \frac{\xi}{4}\right]\right) \geq 0.8.$$

Next, since $\hat{\mu}_t(K_r)$ is a martingale, we have $\mathbb{E}[\hat{\mu}_{T_1}(K_r)] \geq 1 - \zeta/100$. Since $\hat{\mu}_{T_1}(K_r) \leq 1$ almost surely, we can use Markov's inequality to conclude that

$$\mathbb{P}\left(\hat{\mu}_{T_1}(K_r) \leq 1 - \zeta/20\right) \leq 0.2.$$

This concludes the lemma. $\qquad\square$

4.2.2 Analysis of Stage 2

In the second stage of the process, the measure of E is kept constant almost surely, so Lemma 13 boils down to a simple application of Markov's inequality.

Proof of Lemma 13. In the second stage, by construction, we have

$$\hat{\mu}_{T_2}(E) = \hat{\mu}_{T_1}(E).$$

Additionally, since $\hat{\mu}_t(K_r)$ is a martingale, for $t \in [T_1, T_2]$, we have

$$\mathbb{E}[\hat{\mu}_t(K_r) \mid W_1] = \mu_{T_1}(K_r) \geq 1 - \zeta/20.$$

Then $\mathbb{P}(\hat{\mu}_t(K_r) \geq 1 - \zeta/10) + (1 - \zeta/10)\,(1 - \mathbb{P}(\hat{\mu}_t(K_r) \geq 1 - \zeta/10)) \geq 1 - \zeta/20$, which results in

$$\mathbb{P}(\hat{\mu}_t(K_r) \geq 1 - \zeta/10) \geq 0.5.$$

We conclude by applying the union bound. $\qquad\qquad\square$

4.2.3 Analysis of Stage 3

In this section we prove Lemma 14. Here is the main observation: since the driving matrix C_t is fixed to be the rank-1 matrix $\theta\theta^T$ between time T_2 and T_3, the SL process during that time only depends on the marginal of the measure $\hat{\mu}_{T_2}$ onto the direction θ, which means that the proof boils down to the analysis of a one-dimensional SL process. This analysis, however, is quite long and technical and requires a few lemmas on properties of one-dimensional logconcave measures summarized in Appendix A.

Recall that we condition on the event W_2 which amounts to

$$\hat{\mu}_{T_2}(E) \in \left[\xi - \frac{\xi}{4}, \xi + \frac{\xi}{4}\right] \text{ and } \hat{\mu}_{T_2}(K_r) \geq 1 - \zeta/10,$$

For a unit vector θ (obtained by Stage 2) the third stage of the process runs the SL process with a control matrix $C_t = \theta\theta^\top$ time period $\alpha := n - \lambda_1$. That is, for $t \geq T_2$,

$$d\hat{\mu}_t(x) = (x - \mathbf{b}(\hat{\mu}_t))^\top \theta\theta^\top dW_t \hat{\mu}_t(x).$$

Define

$$\sigma^2 := \theta^\top A_{T_2}\theta = \theta^\top \mathrm{Cov}(\hat{\mu}_{T_2})\theta,$$

the variance of the starting measure in the direction of θ. As a result of the Brascamp and Lieb inequality [BL02], we have $A_{T_2} \preceq B_{T_2}^{-1}$, and hence

$$\sigma^2 \leq \frac{1}{T_1}. \tag{16}$$

For $t \in [0, \alpha]$, define ω_t to be the density on $\mathbb{R}$ obtained by taking the push-forward of $\hat{\mu}_{T_2+t}$ via $x \mapsto \frac{1}{\sigma} \cdot x^\top \theta$. That is,

$$\omega_t(z) := \int_{H(z)} \hat{\mu}_{T_2+t}(x)dx, \tag{17}$$

where, for $z \in \mathbb{R}$, $H(z) := \{x \in \mathbb{R}^n;\ \theta^T x = z\sigma\}$ is defined as the fiber corresponding to the value z. For a subset $S \subseteq \mathbb{R}^n$, define

$$h_S(z) := \begin{cases} \dfrac{\int_{H(z)} \mathbf{1}_{\{x \in S\}}\hat{\mu}_{T_2}(x)dx}{\int_{H(z)} \hat{\mu}_{T_2}(x)dx} & \text{if } \omega_0(z) \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Observe that $h_S(z) \in [0, 1]$ for all z. Note that the value of h_S remains unchanged if the term $\hat{\mu}_{T_2}$ is replaced with $\hat{\mu}_{T_2+t}$ in the above formula, for any $t \in [0, \alpha]$. This is because

$$\hat{\mu}_{T_2+t}(x) \propto \hat{\mu}_{T_2}(x) \exp\left(-t(\theta^\top x)^2 + \tilde{c}_t \theta^\top x\right),$$

for some $\tilde{c}_t$. This means that for a fixed value of z the points in $H(z)$ are all multiplied by the same factor. The definition of h_S allows us to identify $\omega_t(h_S)$ with $\hat{\mu}_{T_2+t}(S)$ as shown in the following lemma.

14

Lemma 16. *For the stochastic process ω_t defined above, for $t > 0$, we have*

$$\mathbf{b}(\omega_t) = \theta^\top \mathbf{b}(\hat{\mu}_{T_2+t})/\sigma$$
$$\mathrm{Var}(\omega_t) = \theta^\top \mathrm{Cov}(\hat{\mu}_{T_2+t})\theta/\sigma^2$$
$$\omega_t(h_S) = \hat{\mu}_{T_2+t}(S).$$

In particular, $\mathrm{Var}(\omega_0) = 1$. Additionally, $(\omega_t)_{t\geq 0}$ satisfies the following stochastic differential equation

$$d\omega_t = \sigma(z - \mathbf{b}(\omega_t))d\bar{W}_t\omega_t(z), \tag{19}$$

where $\bar{W}_t = \theta^\top W_t$ is a one-dimensional Brownian motion.

The proof of this lemma is straightforward, and is provided in Subsection 4.2.5 below. Based on the above stochastic differential equation, for $t > 0$, define

$$\mathbf{y}_t := \frac{1}{t}\int_0^t \frac{1}{\sigma}d\bar{W}_s + \mathbf{b}(\omega_s)ds.$$

According to Lemma 2.1 in [Eld13], ω_t is uniquely determined by the value of $\mathbf{y}_t$, for all $t \in [0, \alpha]$. Additionally, it takes the form

$$\omega_t(z) = \frac{\omega_0(z)\exp(-t\sigma^2(z - \mathbf{y}_t)^2)}{\int \omega_0(\varsigma)\exp(-t\sigma^2(\varsigma - \mathbf{y}_t)^2)d\varsigma}.$$

An application of Theorem 2 in [EAM22] shows that $\mathbf{y}_\alpha$ has the law ρ, where

$$\rho := \omega * \mathcal{N}\left(0, \frac{1}{\alpha\sigma^2}\right). \tag{20}$$

With a slight abuse of notation, we introduce a deterministic density $\omega_{y,t}$ with two subindices as

$$\omega_{y,t}(z) := \frac{\omega_0(z)\exp\left(-t\sigma^2(z - y)^2\right)}{\int \omega_0(\varsigma)\exp\left(-t\sigma^2(\varsigma - y)^2\right)d\varsigma}, \quad \forall z \in \mathbb{R}. \tag{21}$$

From here, we can identify ω_t with $\omega_{\mathbf{y}_t,t}$. The next lemma shows that the function $y \mapsto \omega_{y,\alpha}(h_S)$ is Lipschitz for a fixed $S \subset \mathbb{R}$.

Lemma 17. *Let $h : \mathbb{R}^n \to [0, 1]$. Then for all $y, \tilde{y} \in \mathbb{R}$, we have*

$$|\omega_{y,\alpha}(h) - \omega_{\tilde{y},\alpha}(h)| \leq \sqrt{\alpha\sigma^2}|y - \tilde{y}|.$$

Its proof is provided in Subsection 4.2.5. With the above notation and lemmas, we are ready to prove Lemma 14. It is done by discussing the two cases based on the value of $\alpha\sigma^2$:

1. $\alpha\sigma^2 < \frac{1}{512}$
2. $\alpha\sigma^2 \geq \frac{1}{512}$.

Define the events

$$H_1 := \left\{\omega_{Y,\alpha}(h_E) \in \left[\frac{1}{2}\xi, \frac{3}{4}\right]\right\} \cap \left\{\omega_{Y,\alpha}(h_{K_r}) \geq \frac{7}{8}\right\}, \text{ and}$$

$$H_0 := \left\{\omega_0(h_E) \in \left[\frac{3}{4}\xi, \frac{5}{4}\xi\right]\right\} \cap \left\{\omega_0(h_{K_r}) \geq 1 - \frac{\varsigma}{10}\right\}.$$

The results in the two cases are summarized in the two lemmas below.

Lemma 18. *If $\alpha\sigma^2 \leq \frac{1}{512}$ and $\zeta \leq 0.1\xi$, then*

$$\mathbb{P}_{Y\sim\rho}\left(H_1 \mid H_0\right) \geq 0.1\xi.$$

Lemma 19. *If $\alpha\sigma^2 \geq \frac{1}{512}$, $\zeta \leq \xi^2/\sqrt{\log(10^4/\xi)}/10^8$, then there exists a universal constant $C > 0$ such that*

$$\mathbb{P}_{Y\sim\rho}\left(H_1 | H_0\right) \geq \frac{C}{\log(\frac{1}{\xi})}\sqrt{\frac{T_1}{n}}\xi.$$

Proof of Lemma 14. Using the identification in Lemma 16 between the terms in ω and those in $\hat{\mu}$, it is clear that Lemma 14 follows from the two lemmas above. □

4.2.4 Proof of Lemma 18 and Lemma 19

First, we prove Lemma 18 where $\alpha\sigma^2 \leq \frac{1}{512}$. The bound is proven by direct analysis of the process $\omega_t(h_E)$ via the stochastic differential equation (19).

Proof of Lemma 18. Let $g_t := \omega_t(h_E)$. According to Eq. (19), g_t satisfies

$$dg_t = \int_{\mathbb{R}} \sigma(z - \mathbf{b}(\omega_t)) \cdot d\bar{W}_t h_E(z)\omega_t(z)dz.$$

It is a martingale for $t \geq 0$, with $g_0 = w_0(h_E) = \hat{\mu}_{T_2}(E) \in [\frac{3}{4}\xi, \frac{5}{4}\xi]$. We first claim that for any $t_1 \in [0, \alpha]$ to α, we have almost surely that

$$g_{t_1} \in [1/2, 5/8] \quad \Rightarrow \quad \mathbb{P}(g_\alpha \in [3/8, 6/8] \mid \omega_{t_1}) \geq 0.4. \tag{22}$$

The proof of the claim (22) is deferred to the end. Assuming the claim for now, we complete the proof of Lemma 18.

If $g_0 \in [1/2, 5/8]$, then we directly apply the claim (22) from time 0 to α to obtain

$$\mathbb{P}(g_\alpha \in [3/8, 6/8]) \geq 0.4.$$

Otherwise, $g_0 \in [\frac{3}{4}\xi, 1/2)$, we define the stopping time

$$\boldsymbol{\tau} := \min\left\{t \in [0, \alpha] \mid g_t = \frac{1}{2}\xi \text{ or } g_t = \frac{1}{2}\right\} \wedge \alpha. \tag{23}$$

Since $(g_t)_{t\geq 0}$ is a martingale with $\mathbb{E}[g_t] = g_0 \in [\frac{3}{4}\xi, 1/2)$, applying the optimal stopping time theorem for bounded martingale, we have

$$\mathbb{E}[g_{\boldsymbol{\tau}}] = g_0 \in [\frac{3}{4}\xi, 1/2).$$

Separating the above expectation into three cases $g_{\boldsymbol{\tau}} = \frac{1}{2}\xi$, $g_{\boldsymbol{\tau}} = \frac{1}{2}$ or $\boldsymbol{\tau} = \alpha$, we obtain

$$\mathbb{P}\left(g_{\boldsymbol{\tau}} = \frac{1}{2} \text{ or } \boldsymbol{\tau} = \alpha\right) \geq \frac{1}{2}\xi.$$

Under the above event, if $\boldsymbol{\tau} = \alpha$, we have $g_\alpha \in (\frac{1}{2}\xi, \frac{1}{2})$. Otherwise, we apply the claim (22) from time $\boldsymbol{\tau}$ to α to obtain

$$\mathbb{P}(g_\alpha \in [3/8, 6/8]) = \mathbb{P}(g_\alpha \in [3/8, 6/8] \mid g_{\boldsymbol{\tau}} = \frac{1}{2} \text{ or } \boldsymbol{\tau} = \alpha)\mathbb{P}(g_{\boldsymbol{\tau}} = \frac{1}{2} \text{ or } \boldsymbol{\tau} = \alpha)$$

$$\geq 0.4\mathbb{P}(g_{\boldsymbol{\tau}} = \frac{1}{2} \text{ or } \boldsymbol{\tau} = \alpha)$$

$$\geq 0.2\xi.$$

Combining all the cases above, we conclude that $g_\alpha \in [\frac{1}{2}\xi, \frac{3}{4}]$ given $g_0 \in [\frac{3}{4}\xi, \frac{5}{4}\xi]$ with probability at least 0.2ξ. On the other hand, since $(\omega_t(h_{K_r}))_{t\geq 0}$ is a martingale, we have $\mathbb{E}[\omega_\alpha(h_{K_r})|H_0] \geq 1 - \zeta/10$. Thus, by Markov's inequality and by the fact that $\omega_\alpha(h_{K_r}) \leq 1$ almost surely,

$$\mathbb{P}(\omega_\alpha(h_{K_r}) \geq 7/8) \geq 1 - \zeta.$$

Recalling that $\zeta \leq 0.1\xi$, Lemma 18 follows from a union bound.

Proof of the claim (22): The quadratic variation of g_t is

$$d[g]_t = \left| \int \sigma(z - \mathbf{b}(\omega_t)) h_E(z) \omega_t(z) dz \right|^2 dt$$
$$\overset{(i)}{\leq} \int \sigma^2(z - \mathbf{b}(\omega_t))^2 \omega_t(z) dz \cdot \int h_E(z)^2 \omega_t(z) dz dt$$
$$\leq \sigma^2 \mathrm{Var}(\omega_t) dt.$$

(i) follows from the Cauchy-Schwarz inequality. To control $\varphi_t := \mathrm{Var}(\omega_t)$, we observe that it satisfies

$$d\varphi_t = \int_\mathbb{R} \sigma(z - \mathbf{b}(\omega_t))^3 d\bar{W}_t \omega_t(z) dz - \sigma^2 \varphi_t^2 dt.$$

We have $\mathbb{E}d\varphi_t \leq 0$, hence $\mathbb{E}[\varphi_t] \leq \varphi_0 \leq 1$.
Consequently, $\mathbb{E}[g]_\alpha \leq \frac{1}{512}$. Conditioned on $g_0 \in [\frac{1}{2}, \frac{5}{8}]$, we have

$$\mathbb{P}\left(\frac{3}{8} \leq g_\alpha \leq \frac{6}{8} \right) \geq \mathbb{P}\left(-\frac{1}{8} \leq \tilde{W}_{[g]_\alpha} \leq \frac{1}{8} \right)$$
$$= 1 - \mathbb{P}\left(\max_{0 \leq \ell \leq \frac{1}{256}} \left| \tilde{W}_\ell \right| > \frac{1}{8} \right) - \mathbb{P}\left([g]_\alpha > \frac{1}{256} \right)$$
$$\overset{(i)}{\geq} 1 - 4\mathbb{P}(\tilde{W}_{\frac{1}{256}} > \frac{1}{8}) - \mathbb{P}([g]_\alpha > \frac{1}{256})$$
$$\overset{(ii)}{\geq} 0.9 - 0.5 = 0.4$$

(i) follows from the reflection principle. (ii) follows from the tail bound of the normal distribution $\mathbb{P}_{X \sim \mathcal{N}(0,1)}(X > 2) \leq 0.023$ and the Markov inequality. $\qquad \square$

Next, we prove Lemma 19 which concerns the case that $\frac{1}{512} \leq \alpha\sigma^2$. The main idea in the proof is to use the Lipschitz bound established in Lemma 17: Since the function $y \to \omega_{y,\alpha}(h_E)$ is Lipschitz and since it must attain values both close to 0 and to 1, we conclude that there exists an interval with non-negligible mass, on which $\omega_{y,\alpha}(E)$ both is bounded away from both 0 and 1. The main caveat is then to show that $\omega_{y,\alpha}(h_{K_r})$ can be made large at the same time. This crucially relies on the convexity of K_r, used in the following lemma, whose proof is found in the next subsection.

Lemma 20. *Suppose that (i)* $\frac{1}{512} \leq \alpha\sigma^2$*, (ii)* $\zeta \leq \xi^2/\sqrt{\log(10^4/\xi)}/10^8$ *and (iii)* $\omega_0(h_{K_r}) \geq 1 - \zeta/10$*. Then, there exists an interval* $I \subset \mathbb{R}$ *such that* $\rho(I) \geq 1 - \xi/10$ *and of length at most* $100(4 + \log \frac{1}{\xi})$*, such that*

$$\omega_{y,\alpha}(h_{K_r}) \geq 7/8, \quad \forall y \in I.$$

Proof of Lemma 19. Let I be the interval obtained by applying Lemma 20. We partition I into three subsets

$$
\begin{aligned}
I_1 &:= \{y \in \mathbb{R} \mid \omega_{y,\alpha}(h_E) \leq 1/2\xi\} \cap I, \\
I_2 &:= \{y \in \mathbb{R} \mid \omega_{y,\alpha}(h_E) \in [1/2\xi, 3/4]\} \cap I, \\
I_3 &:= \{y \in \mathbb{R} \mid \omega_{y,\alpha}(h_E) \geq 3/4\} \cap I.
\end{aligned}
$$

Since $\mathbb{E}[\omega_{Y,\alpha}(h_E)|H_0] = \omega_0(h_E) \in [3/4\xi, 5/4\xi]$, Markov's inequality implies that

$$
5/4\xi \geq 0 + 3/4 \cdot \mathbb{P}(\omega_{Y,\alpha}(h_E) > 3/4|H_0).
$$

Recalling that the measure ρ defined in Eq. (20) satisfies $\rho(I) \geq 1 - \frac{1}{10}\xi$, we obtain by a union bound that

$$
\mathbb{P}(Y \in I_1 \cup I_2|H_0) = \mathbb{P}(\omega_{Y,\alpha}(h_E) \leq 3/4|H_0) - \frac{1}{10}\xi \geq 1 - \frac{5}{3}\xi - \frac{1}{10}\xi \geq 1/10,
$$

where Y is ρ-distributed. If $\mathbb{P}(Y \in I_2|H_0) \geq 1/20$, then we are done. Otherwise, we obtain

$$
\mathbb{P}(Y \in I_1) \geq 1/20. \tag{24}
$$

Similarly, since

$$
3/4\xi \leq \mathbb{E}[\omega_{Y,\alpha}(h_E)] \leq 1 \cdot \mathbb{P}(\omega_{Y,\alpha}(h_E) \geq 1/2\xi) + 1/2\xi \cdot \mathbb{P}(\omega_{Y,\alpha}(h_E) \leq 1/2\xi),
$$

we obtain

$$
\mathbb{P}(Y \in I_2 \cup I_3) = \mathbb{P}(\omega_{Y,\alpha}(h_E) \geq 1/2\xi) - \frac{1}{10}\xi \geq \frac{1}{10}\xi.
$$

If $\mathbb{P}(Y \in I_2) \geq \frac{1}{20}\xi$, then we are done. Otherwise, we obtain

$$
\mathbb{P}(Y \in I_3) \geq \frac{1}{20}\xi. \tag{25}
$$

Observe that the distance between I_1 and I_3 is large, because for any $y \in I_1$ and $\tilde{y} \in I_3$, we have

$$
\omega_{\tilde{y},\alpha}(h_E) - \omega_{y,\alpha}(h_E) \geq \frac{3}{4} - \frac{1}{2}\xi \geq \frac{1}{2}.
$$

And according to Lemma 17, $y \mapsto \omega_{y,\alpha}(h_E)$ is $\sqrt{\alpha\sigma^2} \leq \sqrt{n/T_1}$-Lipschitz as $\alpha \leq n$ and $\sigma^2 \leq \frac{1}{T_1}$ in Eq. (16). Hence,

$$
|\tilde{y} - y| \geq \frac{1}{2}\sqrt{\frac{T_1}{n}}. \tag{26}
$$

Let $(\mathbf{1}_I \cdot \omega)$ be the measure ω restricted on the set I. Finally, we consider $(\mathbf{1}_I \cdot \omega) * \mathcal{N}\left(0, \frac{1}{\alpha\sigma^2}\right)$, which satisfies a diameter isoperimetric inequality (see Theorem 4.2 in [Vem05]). Hence,

$$
\begin{aligned}
\rho(I_2) &\geq \frac{2d(I_1, I_3)}{100\left(1 + \log(\frac{1}{\xi})\right)} \min\{\rho(I_1), \rho(I_3)\} \\
&\geq \frac{C}{\log(\frac{1}{\xi})}\sqrt{\frac{T_1}{n}}\xi,
\end{aligned}
$$

where the inequality follows from Eq. (24) (25) and (26). $\qquad\square$

4.2.5 Additional proofs in the third SL stage

In this subsection, we complete the missing proofs of the lemmas in the previous subsection.

Proof of Lemma 16. Based on the ω_t definition (17), we can express the mean and variance of ω_t using $\hat{\mu}_{T_2+t}$ as follows.

$$
\begin{aligned}
\mathbf{b}(\omega_t) &= \int z\omega_t(z)dz \\
&= \int \int_{\theta^\top x/\sigma=z} \theta^\top x/\sigma \hat{\mu}_{T_2+t}(x)dxdz \\
&= \theta^\top \mathbf{b}(\hat{\mu}_{T_2+t})/\sigma \\
\mathrm{Var}(\omega_t) &= \int \left(z - \mathbf{b}(\omega_t)\right)\left(z - \mathbf{b}(\omega_t)\right)\omega_t(z)dz \\
&= \int \int_{\sigma\theta^\top x=z} \frac{1}{\sigma^2}\theta^\top (x - \mathbf{b}(\hat{\mu}_t))(x - \mathbf{b}(\hat{\mu}_t))^\top \theta \hat{\mu}_{T_2+t}(x)dxdz \\
&= \theta^\top \mathrm{Cov}(\hat{\mu}_{T_2+t})\theta/\sigma^2.
\end{aligned}
$$

We deduce that $\mathrm{Var}(\omega_0) = 1$ as $\sigma^2 = \theta^\top \mathrm{Cov}(\hat{\mu}_{T_2+t})\theta$. For any $z \in \mathbb{R}$, we have

$$
\begin{aligned}
d\omega_t(z) &= \int_{\theta^\top x/\sigma=z} d\hat{\mu}_{T_2+t}(x)dx \\
&= \int_{\theta^\top x/\sigma=z} (x - \mathbf{b}(\hat{\mu}_{T_2+t}))^\top \theta\theta^\top dW_t \hat{\mu}_{T_2+t}(x)dx \\
&= \int_{\theta^\top x/\sigma=z} \sigma(z - \mathbf{b}(\omega_t))\theta^\top dW_t \hat{\mu}_{T_2+t}(x)dx \\
&= \sigma(z - \mathbf{b}(\omega_t))\theta^\top dW_t \omega_t(z).
\end{aligned}
$$

Based on the above calculation, $(\omega_t)_{t\geq 0}$ undergoes a 1-dimensional SL and the amount of Gaussian that it makes appear at time t is $t\sigma^2$.

Based on the h_S definition, we can express $\hat{\mu}_{T_2+t}(S)$ using ω_t and h_S as follows

$$
\begin{aligned}
\omega_t(h_S) &= \int h_S(z)\omega_t(z)dz \\
&= \int \frac{\int_{\theta^\top x/\sigma=z} \mathbf{1}_{x\in S}\hat{\mu}_{T_2+t}(x)dx}{\omega_t(z)}\omega_t(z)dz \\
&= \hat{\mu}_{T_2+t}(S).
\end{aligned}
$$

$\square$

Proof of Lemma 17. First, we look at the derivative of $\omega_{y,\alpha}$ with respect to y. Recall that

$$
\omega_{y,\alpha} = \frac{\exp(-\frac{\alpha\sigma^2}{2}|x-y|^2)\nu(x)}{\int \exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)\nu(z)dz}.
$$

Taking derivative with respect to y, we obtain

$$
\begin{aligned}
\frac{\partial \omega_{y,\alpha}(x)}{\partial y} &= -\alpha\sigma^2 \left[(y-x) - \frac{\int(y-z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)\omega(z)dz}{\int \exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)\omega(z)dz}\right]\omega_{y,\alpha} \\
&= \alpha\sigma^2[x - \mathbf{b}(\omega_{y,\alpha})]\omega_{y,\alpha}(x).
\end{aligned}
$$

19

Consequently,

$$\frac{\partial \omega_{y,\alpha}(h)}{\partial y} = \int h(x) \cdot \alpha \sigma^2 [x - \mathbf{b}(\omega_{y,\alpha})] \omega_{y,\alpha}(x) dx.$$

Second, let v be the unit vector in the direction of $y - \tilde{y}$. Then $y - \tilde{y} = |y - \tilde{y}| \, v$. Applying the mean value theorem, there exists $\hat{y}$ such that

$$
\begin{aligned}
|\omega_{y,\alpha}(h) - \omega_{\tilde{y},\alpha}(h)| &= \left| (y - \tilde{y})^\top \frac{\partial \omega_{z,\alpha}(h)}{\partial z} |_{z=\hat{y}} \right| \\
&= \alpha \sigma^2 \, |y - \tilde{y}| \left| \int h(x) v \cdot (x - \mathbf{b}(\omega_{\hat{y},\alpha})) \omega_{\hat{y},\alpha}(x) dx \right| \\
&\leq \alpha \sigma^2 \, |y - \tilde{y}| \left[\int (x - \mathbf{b}(\omega_{\hat{y},\alpha}))^2 \, \omega_{\hat{y},\alpha}(x) dx \right]^{1/2} \left[\int h(x)^2 \omega_{\hat{y},\alpha}(x) dx \right]^{1/2} \\
&\leq \alpha \sigma^2 \, |y - \tilde{y}| \, \mathrm{Var}(\omega_{\hat{y},\alpha})^{1/2} \cdot 1 \\
&\leq \alpha \sigma^2 \, |y - \tilde{y}| \, \frac{1}{(\alpha \sigma^2)^{1/2}} \\
&= (\alpha \sigma^2)^{1/2} \, |y - \tilde{y}| \, .
\end{aligned}
$$

The last inequality follows because $\omega_{y,\alpha}$ is $\alpha \sigma^2$-strongly logconcave for any y. This completes the proof. $\square$

Before we can prove Lemma 20, we need the following fact regarding on two logconcave densities.

Lemma 21. *Let $p, \tilde{p}$ be two logconcave densities on $\mathbb{R}$ such that p is isotropic and $\mathrm{d}_{\mathrm{TV}}(p, \tilde{p}) < \epsilon$. Let $0 < \delta < 1$. Set a and b to be the δ and $(1 - \delta)$-quantiles respectively of p, respectively. Suppose further that $\epsilon \leq \delta^2/10^5$. Then for all $z \in [a, b]$, we have $\frac{\tilde{p}}{p}(z) > 0.9$.*

Proof of Lemma 21. Without loss of generality (by convolving with small Gaussian and taking limit), we can assume that $\frac{\tilde{p}}{p}$ is well-defined on $\mathbb{R}$ and both densities are C^2-smooth. Let $\tilde{a}$ and $\tilde{b}$ be the $\delta/2$ and $(1 - \delta/2)$-quantiles of p. Define $f(\cdot) := \frac{\tilde{p}}{p}(\cdot)$.

Suppose there is $z_0 \in [a, b]$ such that $f(z_0) \leq 0.9$. Define $w_- := \sup \{z \in [\tilde{a}, z_0] \mid f(z) \geq 0.95\}$, noting that such $w_- \in [\tilde{a}, z_0]$ must indeed exist, for otherwise we always have $f(x) \leq 0.95$ for $x \in [\tilde{a}, a]$, which implies

$$\epsilon > \mathrm{d}_{\mathrm{TV}}(\nu, \tilde{\nu}) \geq \int_a^{\tilde{a}} (1 - f(x)) p(x) dx \geq \frac{\delta}{2} \cdot 0.05,$$

which contradicts the assumption $\epsilon \leq \delta^2/10^5$ and $\delta < 1$. The same reasoning allows us to define

$$w_+ := \inf \left\{ z \in [z_0, \tilde{b}] \mid f(z) \geq 0.95 \right\}.$$

By the definition of w_- and w_+, we have $f(x) \leq 0.95$ for all $x \in [w_-, w_+]$, which implies that

$$p([w_-, w_+]) \cdot 0.05 \leq \int_{w_-}^{w_+} (1 - f(x)) p(x) dx \leq \mathrm{d}_{\mathrm{TV}}(p, \tilde{p}) < \epsilon.$$

Using Lemma 28 below, we deduce that $p(z) \geq \frac{1}{16e} \delta$ for $z \in [w_-, w_+]$ which implies that

$$|w_+ - w_-| \leq \frac{320e \cdot \epsilon}{\delta}. \tag{27}$$

By the mean value theorem, there exists $u_- \in [w_-, z_0]$ such that

$$
\begin{aligned}
(\log \tilde{p} - \log p)'(u_-) &= \frac{\log f(z_0) - \log f(w_-)}{|z_0 - w_-|} \\
&\leq \frac{-\log \frac{0.95}{0.9}}{|z_0 - w_-|} \\
&\leq \frac{-\log \frac{0.95}{0.9} \cdot \delta}{320 e \cdot \epsilon} =: -\mathfrak{D}.
\end{aligned}
$$

Similarly, there exist u_+ between z_0 and w_+ such that

$$
(\log \tilde{p} - \log p)'(u_+) \geq \mathfrak{D}.
$$

Combining the two bounds above, we have

$$
(\log \tilde{p} - \log p)'(u_+) - (\log \tilde{p} - \log p)'(u_-) \geq 2\mathfrak{D}.
$$

Since $\log \tilde{p}$ is concave, its derivative is non-decreasing, and consequently we have

$$
(\log \tilde{p})'(u_+) - (\log \tilde{p})'(u_-) \leq 0.
$$

It follows that

$$
(\log p)'(u_-) - (\log p)'(u_+) \geq 2\mathfrak{D}.
$$

Then either $(\log p)'(u_-)$ or $-(\log p)'(u_+)$ has to be larger than $\mathfrak{D}$. If $-(\log p)'(u_+) \geq \mathfrak{D}$, because $\log(p)$ is non-increasing, for $z \geq u_+$,

$$
(\log p)'(x) \leq -\mathfrak{D}.
$$

Integrating leads to the bound $p(z) \leq p(u_+) e^{-\mathfrak{D}(z - u_+)}$. Integrating again, we have

$$
\int_{u_+}^{\infty} p(z) dz \leq p(u_+) \frac{1}{\mathfrak{D}} \leq \frac{1}{\mathfrak{D}} = \frac{320 e \cdot \epsilon}{\log \frac{0.95}{0.9} \cdot \delta} \leq \frac{20000 \epsilon}{\delta} \leq \frac{1}{5} \delta,
$$

where $p(u_+) \leq 1$ follows from Lemma 25. This leads to a contradiction to the assumption that the mass on the right of u_+ is at least $\frac{1}{2}\delta$. The other case $(\log p)'(u_-) \geq \mathfrak{D}$ is similar. $\qquad\square$

Proof of Lemma 20. The proof is split into two disjoint cases: $\frac{1}{512} \leq \alpha\sigma^2 < \frac{400 \log(10^4/\xi)}{\xi^2}$ and $\alpha\sigma^2 \geq \frac{400 \log(10^4/\xi)}{\xi^2}$.

Case 1: Assume that $\frac{1}{512} \leq \alpha\sigma^2 < \frac{400 \log(10^4/\xi)}{\xi^2}$.

Let $\delta = \xi/20$. Let a and b be the δ and $(1-\delta)$-quantiles of ρ. In this case, we show that the interval $I = [a, b]$ satisfies the requirements of the lemma. In this case we trivially have $\rho(I) \geq 1 - \xi/10$.

Since ρ is the convolution of ω which has variance 1 and $\mathcal{N}\left(0, \frac{1}{\alpha\sigma^2}\right)$ which has variance 512, the variance of ρ is upper-bounded by 513.

Applying the tail bound for isotropic log-concave density in Lemma 27, we obtain

$$
|a - b| \leq 2\sqrt{513} \log \frac{e}{\delta} \leq 100 \left(4 + \log \frac{1}{\xi}\right).
$$

It remains to show that $\omega_{y,\alpha}(h_{K_r}) \geq 7/8$ for all $y \in I$. Let $J := \{y \in [a,b] \mid \omega_{y,\alpha}(h_{K_r}) \geq \frac{15}{16}\}$. Note that J is not necessarily convex. Since $\mathbb{E}_{Y \sim \rho}\omega_{Y,\alpha}(K_r) = \omega_0(K_r) \geq 1 - \zeta/10$, we have by Markov's inequality that

$$\mathbb{P}_{Y \sim \rho}\left(\omega_{Y,\alpha}(K_r) \geq \frac{15}{16}\right) \geq 1 - 2\zeta. \tag{28}$$

Fix any $z_0 \in [a,b] \setminus J$, define the largest interval around z_0 contained in $[a,b] \setminus J$ as follows

$$z_+ = \sup\{\tilde{z} \mid [z_0, \tilde{z}] \subseteq [a,b] \setminus J\}$$
$$z_- = \inf\{\tilde{z} \mid [\tilde{z}, z_0] \subseteq [a,b] \setminus J\}.$$

For any $\tilde{z} \in [z_-, z_+]$, we have $\omega_{\tilde{z},\alpha}(h_{K_r}) < \frac{15}{16}$ according to the definition of J. It follows from Eq. (28) that $\rho([z_-, z_+]) \leq 2\zeta$. On the other hand, since the variance of ρ is bounded by 513, applying Lemma 28, we obtain $\rho(z) \geq \frac{1}{8\sqrt{513e}}\delta$ for $z \in [a,b]$. Hence,

$$|z_+ - z_-| \leq \frac{2\zeta}{\frac{1}{8\sqrt{513e}}\delta} \leq \frac{2000\zeta}{\delta}.$$

Finally, we can use the Lipschitz property of $y \mapsto \omega_{y,\alpha}(h_{K_r})$ to lower bound $\omega_{\tilde{z},\alpha}(h_{K_r})$ for $\tilde{z} \in [z_-, z_+]$. According to Lemma 17, $y \mapsto \omega_{y,\alpha}(h_{K_r})$ is $\sqrt{\alpha\sigma^2} \leq \sqrt{\frac{400\log(10^4/\xi)}{\xi^2}}$-Lipschitz. Note that the assumption that $\zeta \leq \xi^2/\sqrt{\log(10^4/\xi)}/10^8$ implies $\frac{2000\zeta}{\delta} \cdot \frac{\sqrt{400\log(10^4/\xi)}}{\xi} \leq \frac{1}{16}$, which in turn ensures that

$$\omega_{\tilde{z},\alpha}(h_{K_r}) \geq 7/8, \text{ for } \tilde{z} \in [z_-, z_+].$$

Hence, for any $y \in [a,b]$, we have $\omega_{y,\alpha}(h_{K_r}) \geq 7/8$. This concludes the case $\frac{1}{512} \leq \alpha\sigma^2 \leq \frac{400\log(10^4/\xi)}{\xi^2}$ with the choice of the interval $I = [a,b]$.

Case 2: Assume that $\alpha\sigma^2 \geq \frac{400\log(10^4/\xi)}{\xi^2}$.
Take $\delta = \xi/100$. Let a and b be the δ and $(1-\delta)$-quantiles of ω, respectively. Let $I := [a+\delta, b-\delta]$. Using the same reasoning as in the beginning of the previous case, we know that its length satisfies $|I| \leq |a - b| \leq 100\left(4 + \log\frac{1}{\xi}\right)$.

Next, let us show that $\rho(I) \geq 1 - \xi/10$. To that end, let $X \sim \omega$ and $Z \sim \mathcal{N}(0, \frac{1}{\alpha\sigma^2})$ be independent of each other. Then

$$\begin{aligned}
\rho([a+\delta, b-\delta]) &= \mathbb{P}(X + Z \in [a+\delta, b-\delta]) \\
&\geq \mathbb{P}(X \in [a+2\delta, b-2\delta] \text{ and } Z \in [-\delta, \delta]) \\
&= \mathbb{P}(X \in [a+2\delta, b-2\delta]) \cdot \mathbb{P}(Z \in [-\delta, \delta]) \\
&\overset{(v)}{\geq} (\omega([a,b]) - 4\delta \cdot 1) \cdot (1 - \delta/500) \\
&\geq 1 - 7\delta \geq 1 - \xi/10
\end{aligned}$$

The step (v) follows because $\omega([a,b]) = 1 - 2\delta$, $\omega[a+2\delta, b-2\delta] \geq \omega([a,b]) - 4\delta \cdot 1$ using Lemma 25 and by the Gaussian tail bound in Eq. (30).

To conclude the proof it remains to show that for $y \in I$ one has $\omega_{y,\alpha}(h_{K_r}) \geq 7/8$, where

$$\begin{aligned}
\omega_{y,\alpha}(h_{K_r}) &= \frac{\int \omega(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)h_{K_r}(z)dz}{\int \omega(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz} \\
&= \hat{\mu}_{T_2}(K_r)\frac{\int \tilde{\omega}(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz}{\int \omega(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz},
\end{aligned}$$

and where $\tilde{\omega}(x) = \frac{\omega(x)h_{K_r}(x)}{\hat{\mu}_{T_2}(K_r)}$.

Observe that $\tilde{\omega}$ is, by definition, the push-forward of the measure $\tilde{\mu}_{T_2} := \hat{\mu}_{T_2} \cdot \mathbf{1}_{K_r}/\hat{\mu}_{T_2}(K_r)$ via $x \mapsto \frac{1}{\sigma} \cdot x^\top \theta$. Since by construction $d_{TV}(\hat{\mu}_{T_2}, \tilde{\mu}_{T_2}) \leq 1 - \hat{\mu}_{T_2}(K_r) \leq \zeta/10$, we obtain

$$d_{TV}(\omega, \tilde{\omega}) \leq d_{TV}(\hat{\mu}_{T_2}, \tilde{\mu}_{T_2}) \leq \zeta/10.$$

By the Prékopa-Leindler inequality, we have that $\tilde{\omega}$ is logconcave, thus we may apply Lemma 21 with $\epsilon = \zeta/10 \leq \delta^2/10^5$, we deduce that

$$\frac{\tilde{\omega}}{\omega}(z) \geq 0.9, \quad \forall z \in [a, b]. \tag{29}$$

Recall that the standard Gaussian tail bound for $Z \sim \mathcal{N}(0, \frac{1}{\alpha\sigma^2})$ implies for $\eta > 0$,

$$\mathbb{P}\left(|Z| \geq \frac{\eta}{\sqrt{\alpha\sigma^2}}\right) \leq 2e^{-\eta^2/2}. \tag{30}$$

Take $\eta = 2\log^{1/2}(\frac{1000}{\delta})$, then $e^{-\eta^2/2} \leq \delta/1000$. Note that the assumption $\alpha\sigma^2 \geq \frac{400\log(10^4/\xi)}{\xi^2}$ ensures that $\frac{\eta}{\sqrt{\alpha\sigma^2}} \leq \delta$. We have

$$M_1 := \int_{[y-\delta,y+\delta]} \omega(z)\frac{1}{\sqrt{2\pi\frac{1}{\alpha\sigma^2}}} \exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz$$

$$\stackrel{(i)}{\geq} \frac{1}{8e}\delta \int_{[y-\delta,y+\delta]} \omega(z)\frac{1}{\sqrt{2\pi\frac{1}{\alpha\sigma^2}}} \exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz$$

$$\geq \frac{1}{8e}\delta(1 - 2e^{-\eta^2/2}) \geq \delta/50,$$

where (i) follows from Lemma 28.

$$M_2 := \int_{[y-\delta,y+\delta]^c} \omega(z)\frac{1}{\sqrt{2\pi\frac{1}{\alpha\sigma^2}}} \exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz \tag{31}$$

$$\stackrel{(ii)}{\leq} \int_{[y-\delta,y+\delta]^c} \frac{1}{\sqrt{2\pi\frac{1}{\alpha\sigma^2}}} \exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz \tag{32}$$

$$\leq 2e^{-\eta^2/2} \leq \delta/500, \tag{33}$$

where (ii) follows from Lemma 25. The two above displays imply that

$$\int \omega(z)\exp\left(-\frac{\alpha\sigma^2}{2}|z-y|^2\right)dz \leq \frac{10}{9}\int_{[y-\delta,y+\delta]} \omega(z)\exp\left(-\frac{\alpha\sigma^2}{2}|z-y|^2\right)dz. \tag{34}$$

Hence, for $y \in [a+\delta, b-\delta]$, we have

$$\omega_{y,\alpha}(h_{K_r}) = \hat{\mu}_{T_2}(K_r)\frac{\int \tilde{\omega}(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz}{\int \omega(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz}$$

$$\stackrel{(34)}{\geq} 0.9\hat{\mu}_{T_2}(K_r)\frac{\int_{[y-\delta,y+\delta]} \tilde{\omega}(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz}{\int_{[y-\delta,y+\delta]} \omega(z)\exp(-\frac{\alpha\sigma^2}{2}|z-y|^2)dz}$$

$$\stackrel{(29)}{\geq} 0.9^2\hat{\mu}_{T_2}(K_r) \geq \frac{7}{8}.$$

This completes the proof. $\qquad\square$

4.3 A uniqueness result for the distribution attained by stochastic localization

In this subsection, we prove Lemma 10, which follows from the inversion of the multidimensional Mellin transform [Ant07] (also related to the inverse Laplace transform). To see how the problem is related to the inversion of the multidimensional Mellin transform, we first provide some background.

The Mellin transform [Mel96] (see also [Ant07]) of a function $\Phi(x)$ defined in the positive orthant $\mathbb{R}^n_+$ is given by the integral

$$\mathfrak{M}[\Phi](z) = \int_{\mathbb{R}^n_+} \Phi(x) x^{z-1} dx.$$

Suppose we know that for $t > 0$,

$$\mu(x) = \int_{\mathbb{R}^n} \nu_{y,t}(x) p(y) dy, \forall x \in \mathbb{R}^n,$$

where $\nu_{y,t}(x) = \frac{\exp\left(-\frac{t}{2}|x-y|^2\right)\mu(x)}{\int \exp\left(-\frac{t}{2}|z-y|^2\right)\mu(z)dz}$. Suppose $\tilde{p}$ also satisfy

$$\mu(x) = \int_{\mathbb{R}^n} \nu_{y,t}(x) \tilde{p}(y) dy, \forall x \in \mathbb{R}^n.$$

Taking the difference between the two and rearranging the terms that depending on y and x in the integral separately, we obtain

$$0 = \mu(x) \exp\left(-\frac{t}{2}|x|^2\right) \int_{\mathbb{R}^n} \frac{(p(y) - \tilde{p}(y)) \exp\left(-\frac{t}{2}|y|^2\right)}{\int \exp\left(-\frac{t}{2}|z-y|^2\right)\mu(z)dz} \cdot \exp(tx \cdot y) dy.$$

Using the change of variable $w = \exp(ty) \in \mathbb{R}^n_+$, we obtain

$$0 = \mu(x) \exp\left(-\frac{t}{2}|x|^2\right) \int_{\mathbb{R}^n_+} \frac{\left[p\left(\frac{1}{t}\log w\right) - \tilde{p}\left(\frac{1}{t}\log w\right)\right] \exp\left(-\frac{1}{2t}|\log w|^2\right)}{\int \exp\left(-\frac{t}{2}\left|z-\frac{1}{t}\log w\right|^2\right)\mu(z)dz} \cdot \frac{1}{t} w^{x-1} dw.$$

Since $\mu(x) \neq 0$ for $x \in K$, we have that for $x \in K$,

$$0 = \int_{\mathbb{R}^n_+} \frac{\left[p\left(\frac{1}{t}\log w\right) - \tilde{p}\left(\frac{1}{t}\log w\right)\right] \exp\left(-\frac{1}{2t}|\log w|^2\right)}{\int \exp\left(-\frac{t}{2}\left|z-\frac{1}{t}\log w\right|^2\right)\mu(z)dz} \cdot w^{x-1} dw. \tag{35}$$

It is now clear that $x \mapsto 0$ is the Mellin transform of $w \mapsto \frac{\left[p\left(\frac{1}{t}\log w\right) - \tilde{p}\left(\frac{1}{t}\log w\right)\right] \exp\left(-\frac{1}{2t}|\log w|^2\right)}{\int \exp\left(-\frac{t}{2}\left|z-\frac{1}{t}\log w\right|^2\right)\mu(z)dz}$.

Since $w \mapsto \frac{1}{t}\log w$ is a one-to-one mapping from $\mathbb{R}^n_+$ to $\mathbb{R}^n$, in order to show that p is uniquely defined, it is sufficient to show that the inversion of the multidimensional Mellin transform is well-defined.

Proof of Lemma 10. The main strategy is to verify the conditions and then to apply Theorem 2 in [Ant07] on the inversion of the Mellin transform for the functions appeared in Eq. (35). In the notation of [Ant07], take the convex set $U := K \subset \mathbb{R}^n$ and the convex set $\Theta := \mathbb{B}^n(0, 2\pi)$, and set

$$\mathcal{F}(\cdot) := 0$$

on $U + i\mathbb{R}^n$. $\mathcal{F}$ is holomorphic on $U + i\mathbb{R}^n$ and bounded. Applying Theorem 2 in [Ant07], its inverse Mellin transform is well-defined, and it is 0 almost everywhere. We conclude that the function

$$w \mapsto \frac{\left[p\left(\frac{1}{t}\log w\right) - \tilde{p}\left(\frac{1}{t}\log w\right)\right] \exp\left(-\frac{1}{2t}|\log w|^2\right)}{\int \exp\left(-\frac{t}{2}\left|z-\frac{1}{t}\log w\right|^2\right)\mu(z)dz}$$

is 0 almost everywhere, which implies $p = \tilde{p}$ almost everywhere. $\qquad\square$

5 Conductance for the transformed density

In this section we first prove Lemma 9 and then prove Lemma 8.

5.1 Distance of a typical point from μ_t to its center

The main idea to prove Lemma 9 is to first use the observation in [EAM22] on the distribution of c_n and then apply the standard Gaussian concentration.

Proof of Lemma 9. Recall that μ_n is uniquely defined given c_n. The first part of the lemma is a direct application of Theorem 2 in [EAM22]. Additionally, given the data generation $X \sim \mu, Z \sim \mathcal{N}(0, \frac{1}{n}\mathbb{I}_n)$ independent, and $c_n/n \sim X + Z$, a point drawn from μ_t has the same law as the conditional distribution

$$X \mid c_n.$$

We are interested in the conditional distribution $X - \frac{c_n}{n} \mid c_n$. Applying the standard chi-square tail bound (see Lemma 1 in [LM00]), we obtain the unconditional bound

$$\mathbb{P}(|Z|^2 \geq 2) \leq e^{-\frac{n}{16}}$$
$$\mathbb{P}(|Z|^2 \leq \frac{1}{2}) \leq e^{-\frac{n}{16}}.$$

Hence,

$$\mathbb{P}(|Z| \geq \sqrt{2} \text{ or } |Z| \leq \frac{\sqrt{2}}{2}) \leq 2e^{-\frac{n}{16}}. \tag{36}$$

Let $\mathfrak{E}$ denote the event $\mathfrak{E} := \left\{ a \in \mathbb{R}^n \mid \mathbb{P}(|Z| > \sqrt{2} \text{ or } |Z| \leq \frac{\sqrt{2}}{2} \mid c_n = a) > e^{-\frac{n}{32}} \right\}$. Writing out Eq. (36) with conditional probability, we obtain

$$2e^{-\frac{n}{16}} \geq \mathbb{P}(|Z| \geq \sqrt{2} \text{ or } |Z| \leq \frac{\sqrt{2}}{2}) \geq \mathbb{P}(|Z| \geq \sqrt{2} \text{ or } |Z| \leq \frac{\sqrt{2}}{2} \mid c_n \in \mathfrak{E}) \cdot \mathbb{P}(\mathfrak{E}) > e^{-\frac{n}{32}}\mathbb{P}(\mathfrak{E}).$$

Hence, $\mathbb{P}(\mathfrak{E}) < 2e^{-\frac{n}{32}}$. $\qquad\square$

5.2 Overlap bound: Proof of Lemma 8

To lower bound the s-conductance in Lemma 8, we first need to bound the transition overlap for close points in K_r defined in Eq. (3). This concept of transition overlap was previously proposed in [Lov99]. In order to introduce the transition overlap, we first define the notion of $1/8$-quantile hit-and-run step-size.

Definition 22 ($1/8$-quantile hit-and-run step-size). Given a target density ν supported on a convex set K. For $x \in K$, define $F_x(\nu)$, the median step-size of the hit-and-run with target density ν, as the step-size such that

$$\mathbb{P}_{Y \sim P_{x \to \cdot}(\nu)} (|Y - x| \leq F_x(\nu)) = \frac{1}{8}. \tag{37}$$

The hit-and-run transition kernel $y \mapsto P_{x \to y}$ is continuous on K for any $x \in K$, so the above quantity is well-defined.

Lemma 23. *Fix $\beta \in \mathbb{R}^n$ and the target density $\nu_{\beta,n}$. Let $u, v \in K_r$ such that $|u - \beta| \in \left[\frac{1}{\sqrt{2}}, \sqrt{2}\right]$. Suppose that*

$$|u - v| < \frac{2}{\sqrt{n}} \min \left\{ F_u(\nu_{\beta,n}), 1 \right\},$$

then there exists a universal constant $C \in (0, 1)$ such that

$$d_{TV} (P_{u \to \cdot}(\nu_{\beta,n}), P_{v \to \cdot}(\nu_{\beta,n})) < 1 - C \min \left\{ \sqrt{n} F_u(\nu_{\beta,n}), 1 \right\}.$$

To prove the transition overlap bound in Lemma 23, we need to obtain a rough estimate of F_u.

Lemma 24. *Let $r > 0$. For any $\beta \in \mathbb{R}^n$ and $u \in K_r$ such that $|u - \beta| \leq \sqrt{2}$, we have*

$$F_u(\nu_{\beta,n}) \geq \frac{1}{128} \min \left\{ 2r, \frac{1}{8\sqrt{n}} \right\}.$$

Deferring the proof of Lemma 24 to the end, we are ready to prove Lemma 23.

Proof of Lemma 23. For the sake of brevity, throughout the proof we omit the expression $\nu_{\beta,n}$ in our notation (abbreviating $P = P(\nu_{\beta,n})$ for example), since $\nu_{\beta,n}$ is the only target density considered in this proof. It is sufficient to prove that for any $A \subset K$ nonempty measurable subset, we have

$$P_{u \to A} - P_{v \to A} < 1 - C \min \left\{ \sqrt{n} F_u, 1 \right\}$$

Since $P_{u \to A} \leq 1$, this is implied by showing that

$$P_{v \to A} \geq C \min \left\{ \sqrt{n} F_u, 1 \right\}. \tag{38}$$

We partition A into four parts as follows

- A_1 is the part too close to u

$$A_1 := \left\{ x \in A : |x - u| < F_u \right\}.$$

- A_2 is the part which is not almost orthogonal to $u - v$

$$A_2 := \left\{ x \in A : \left| (x - u)^\top (u - v) \right| > \frac{2}{\sqrt{n}} |x - u| \cdot |u - v| \right\}.$$

- A_3 is the part for which the angle $\angle p_u(x)\beta u$ satisfies $\sin(\angle p_u(x)\beta u) > \frac{2}{\sqrt{n}}$, as illustrated in Figure 5.2

$$A_3 := \left\{ x \in A : \sin\left(\angle p_u(x)\beta u \right) > \frac{2}{\sqrt{n}} \right\},$$

where $p_u(x)$ is the projection of β onto the line through u and x. We omit the dependency on x and use p_u when the dependency on x is clear.

- $S := A \setminus (A_1 \cup A_2 \cup A_3)$ is the rest.

The proof proceeds in two steps:

1. Show that $P_{u \to S} \geq \frac{1}{4}$.

2. Show that there exists a constant $C' > 0$ such that $P_{v \to S} \geq C' \cdot P_{u \to S}$. According to the definition of $P_{u \to \cdot}$ in Eq. (1), we need to

   - show that $|x - v|$ can be upper bounded via $|x - u|$,
   - show that $\nu(\ell_{vx})$ can be upper bounded via $\nu(\ell_{ux})$.

Note that the two steps establish a lower bound on $P_{v \to A}$, which concludes the proof in light of (38).

Step 1. According to the definition (37) of F_u, we have

$$P_{u \to A_1} \leq P_{u \to \mathbb{B}_u^n(F_u)} = \frac{1}{8}.$$

Given u, v, $P_{u \to A_2}$ only depends on the uniform distribution on the unit sphere. We evoke the following well-known result [Tko12, B$^+$97] on the area upper bound of a spherical cap of angle $\phi \in (0, \frac{\pi}{2})$,

$$\frac{\mathcal{A}_n(\phi)}{2\mathcal{A}_n(\pi/2)} \leq e^{-n\cos(\phi)^2/2}, \tag{39}$$

where $\mathcal{A}_n(\phi)$ denotes the area of the cap of angle ϕ of a unit n-sphere (unit sphere in $\mathbb{R}^n$). Applying Eq. (39) with $\cos(\phi) = \frac{2}{\sqrt{n}}$, we have

$$P_{u \to A_2} \leq 0.3.$$

Similarly, given u, $P_{u \to A_3}$ only depends on the uniform distribution on the unit sphere and an application of Eq. (39) implies

$$P_{u \to A_3} \leq 0.3.$$

Combining the three displays above, we obtain

$$P_{u \to S} \geq 1 - \frac{1}{8} - 0.3 - 0.3 \geq \frac{1}{4}.$$

Step 2. First, for $x \in S$, we have

$$|x - u| \overset{(i)}{\geq} F(u) \overset{(ii)}{\geq} \frac{\sqrt{n}}{2} |u - v|. \tag{40}$$

(i) follows from $x \notin A_1$. (ii) follows from the assumption of the lemma. Second, we show that $|x - v|$ can be upper bounded via $|x - u|$ as follows

$$
\begin{aligned}
|x - v|^2 &= |x - u|^2 + |u - v|^2 + 2(x - u)^\top (u - v) \\
&\overset{(i)}{\leq} |x - u|^2 + |u - v|^2 + \frac{4}{\sqrt{n}} |x - u| \, |u - v| \\
&\overset{(ii)}{\leq} |x - u|^2 + \frac{4}{n} |x - u|^2 + \frac{8}{n} |x - u|^2 \\
&\leq \left(1 + \frac{12}{n}\right) |x - u|^2 \\
&\leq \left(\left(1 + \frac{6}{n}\right) |x - u|\right)^2.
\end{aligned} \tag{41}
$$

(i) follows from $x \notin A_2$. (ii) follows from Eq. (40). Third, let q_u and q_v be the unit vectors parallel to $u - x$ and $v - x$ respectively, we have

$$
\begin{aligned}
\frac{\nu(\ell_{vx})}{\nu(\ell_{ux})} &= \frac{\int_{p_v + tq_v \in K} \nu(p_v + tq_v) dt}{\int_{p_u + tq_u \in K} \nu(p_u + tq_u) dt} \\
&= \frac{\int_{p_v + tq_v \in K} e^{-\frac{n}{2}|p_v - \beta|^2 - \frac{n}{2}t^2} dt}{\int_{p_u + tq_u \in K} e^{-\frac{n}{2}|p_u - \beta|^2 - \frac{n}{2}t^2} dt} \\
&= \underbrace{\frac{e^{-\frac{n}{2}|p_v - \beta|^2}}{e^{-\frac{n}{2}|p_u - \beta|^2}}}_{Q_1(x)} \underbrace{\frac{\int_{p_v + tq_v \in K} e^{-\frac{n}{2}t^2} dt}{\int_{p_u + tq_u \in K} e^{-\frac{n}{2}t^2} dt}}_{Q_2(x)}.
\end{aligned} \tag{42}
$$

27

To derive an upper bound for the ratio $Q_2(x)$ we, roughly speaking, use the fact that the chord $\ell_{ux} \cap K$ contains an interval, centered at p_u, whose length is larger than F_u. From the definition of A_3, we have $|p_u - u| = \sin(\angle p_u(x)\beta u)\,|u - \beta| \leq \frac{2}{\sqrt{n}}|u - \beta| \leq \frac{2\sqrt{2}}{\sqrt{n}}$. From the definition of A_1, we have $|x - u| \geq F_u$. If $F_u \leq \frac{2}{\sqrt{n}}$, then $\frac{1}{\sqrt{2\pi/n}}\int_{p_u + tq_u \in K} e^{-\frac{n}{2}t^2}dt \geq \sqrt{n}F_u \cdot \gamma(2\sqrt{2} + 2)$; otherwise, $\frac{1}{\sqrt{2\pi/n}}\int_{p_u + tq_u \in K} e^{-\frac{n}{2}t^2}dt \geq \Gamma(2\sqrt{2} + 2) - \Gamma(2\sqrt{2})$.

Combining both cases, we conclude that R_2 obeys the bound

$$Q_2(x) = \frac{\frac{1}{\sqrt{2\pi/n}}\int_{p_v + tq_v \in K} e^{-\frac{n}{2}t^2}dt}{\frac{1}{\sqrt{2\pi/n}}\int_{p_u + tq_u \in K} e^{-\frac{n}{2}t^2}dt} \leq \frac{1}{\min\left\{\sqrt{n}F_u \cdot \gamma(2\sqrt{2} + 2), \Gamma(2\sqrt{2} + 2) - \Gamma(2\sqrt{2})\right\}}. \quad (43)$$

To upper bound the first ratio $Q_1(x)$, it is sufficient to upper bound $-|p_v - \beta|^2 + |p_u - \beta|^2$. Let $\delta_1 := \angle u\beta v, \delta_2 := \angle uxv$. According to the definition of A_1, we have

$$\sin(\delta_2) \leq \frac{|u - v|}{|u - x|} \leq \frac{2}{\sqrt{n}}. \quad (44)$$

From the assumption, we have

$$|u - v| \leq \frac{2}{\sqrt{n}}, \quad (45)$$

and $|u - \beta| \geq \frac{\sqrt{2}}{2}$. Hence,

$$\sin(\delta_1) \leq \frac{|u - v|}{|u - \beta|} \leq \frac{2\sqrt{2}}{n}. \quad (46)$$

To lower-bound $|p_v - \beta|$, we need to upper-bound $|p_v - v|$ and hence the angle $\angle v\beta p_v$. Note that
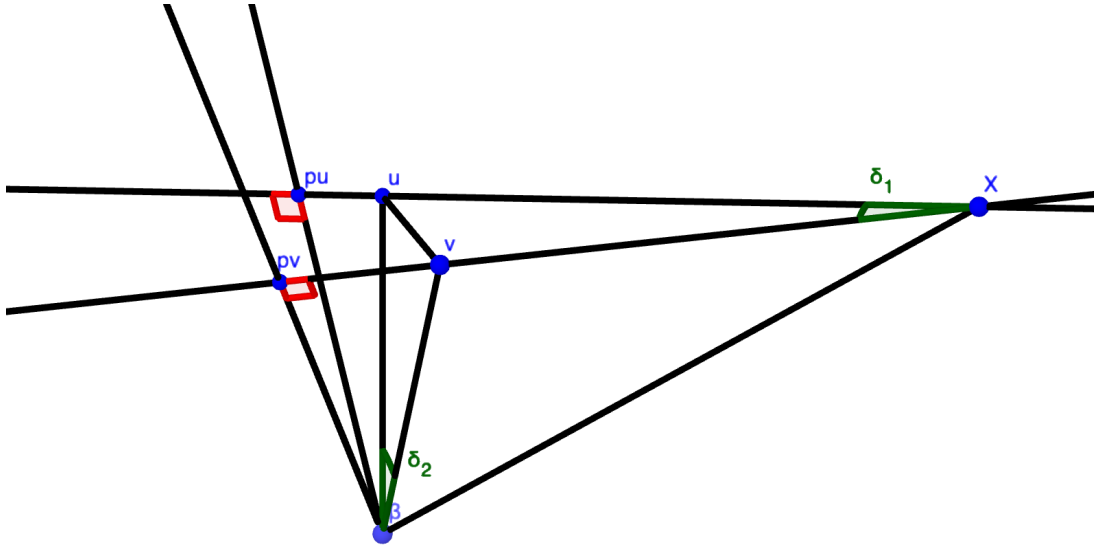

Figure 1: u, v, β, x placed in a 3D plot

$\frac{\pi}{2} - \angle v\beta p_v = \angle p_v v\beta = \angle v\beta x + \angle vx\beta$. The two angles $\angle v\beta x$ and $\angle vx\beta$ can be bounded via $\angle u\beta x$ and $\angle ux\beta$ respectively.

Let $\Delta_1 = \angle u\beta x - \angle v\beta x$ and $\Delta_2 = \angle ux\beta - \angle vx\beta$. We claim that $|\Delta_i| \leq \delta_i$ for $i = 1, 2$. Indeed, if $w_1 = \frac{u - \beta}{|u - \beta|}$, $w_2 = \frac{v - \beta}{|v - \beta|}$ and $w_3 = \frac{x - \beta}{|x - \beta|}$ then by the spherical triangle inequality we have

$$|\arccos(\langle w_1, w_3\rangle) - \arccos(\langle w_2, w_3\rangle)| \leq \arccos(\langle w_1, w_2\rangle),$$

28

which amounts to the fact that $|\Delta_1| \leq \delta_1$, and an analogous argument shows that $|\Delta_2| \leq \delta_2$. We therefore have

$$
\begin{aligned}
|p_v - v| &= |v - \beta| \sin(\angle p_v \beta v) \\
&= |v - \beta| \sin(\frac{\pi}{2} - \angle v \beta x - \angle v x \beta) \\
&= |v - \beta| \sin(\frac{\pi}{2} - \angle u \beta x - \angle u x \beta + \Delta_1 + \Delta_2) \\
&= |v - \beta| \sin(\angle p_u \beta u + \Delta_1 + \Delta_2) \\
&= |v - \beta| \left(\sin(\angle p_u \beta u) + \sin(|\Delta_1| + \sin(|\Delta_2|)\right) \\
&\leq |v - \beta| \left(\frac{|p_u - u|}{|u - \beta|} + |\sin(\delta_1)| + |\sin(\delta_2)|\right).
\end{aligned}
$$

Plugging the bounds (44) (46) into the above equation, together with the definition of A_3, we obtain

$$
\frac{|p_v - v|}{|v - \beta|} \leq \frac{|p_u - u|}{|u - \beta|} + \frac{2\sqrt{2} + 2}{\sqrt{n}} \leq \frac{10}{\sqrt{n}}.
$$

Together with the assumption $|v - \beta| \leq \sqrt{2}$, we deduce that

$$
|p_v - v| \leq \frac{10\sqrt{2}}{\sqrt{n}}. \tag{47}
$$

Using the fact that p_u and p_v are orthogonal projections, we have

$$
\begin{aligned}
-|p_v - \beta|^2 + |p_u - \beta|^2 &= -|v - \beta|^2 + |p_v - v|^2 + |u - \beta|^2 - |p_u - u|^2 \\
&\leq |u - v| \left(2|u - \beta| + |u - v|\right) + |p_v - v|^2 - |p_u - u|^2 \\
&\leq |u - v| \left(2|u - \beta| + |u - v|\right) + |p_v - v|^2 \\
&\overset{(45)\wedge(47)}{\leq} \frac{4}{n}\left(2\sqrt{2} + \frac{4}{n}\right) + \frac{200}{n} \\
&\leq \frac{230}{n},
\end{aligned}
$$

where the first inequality is obtained by the triangle inequality. We therefore have

$$
Q_1(x) = \frac{e^{-\frac{n}{2}|p_v - \beta|^2}}{e^{-\frac{n}{2}|p_u - \beta|^2}} \leq e^{115}. \tag{48}
$$

Combining the bounds for the two ratios, we obtain

$$
\begin{aligned}
P_{v \to S} &= \frac{2}{n\pi_n} \int_S \frac{f(x)}{\nu(\ell_{vx})|x - v|^{n-1}} dx \\
&\overset{(i)}{\geq} \frac{2}{n\pi_n} \int_S \frac{f(x)}{Q_1(x)Q_2(x)\nu(\ell_{ux})|x - v|^{n-1}} dx \\
&\overset{(ii)}{\geq} \frac{2}{e^6 \cdot n\pi_n} \int_S \frac{f(x)}{Q_1(x)Q_2(x)\nu(\ell_{ux})|x - u|^{n-1}} dx \\
&\overset{(43)\wedge(48)}{\geq} C \min\left\{\sqrt{n}F_u, 1\right\} \cdot P_{u \to S},
\end{aligned}
$$

where (i) applies Eq. (42). (ii) follows from (41) and $(1 + \frac{6}{n})^n \leq e^6$. Here C is a universal constant that depends on the universal constant that appear in Eq. (43) and (48).

Finally, we have

$$P_{u \to A} - P_{v \to A} \leq 1 - P_{v \to S}$$
$$\leq 1 - C \min\left\{\sqrt{n} F_u, 1\right\} \cdot P_{u \to S}$$
$$\leq 1 - \frac{1}{4} C \min\left\{\sqrt{n} F_u, 1\right\}.$$

We conclude. Remark that the constants were not optimized for the sake of the simplicity of the derivation. □

Proof of Lemma 24. The lower bound proof proceeds similarly as that of Lemma 3.2 in [LV06b], except that we have to deal with the Gaussian supported on the convex set $\nu_{\beta,n}$. Define $s : K \to \mathbb{R}^+$ as

$$s(u) := \sup\left\{t \in \mathbb{R}_+ \,\middle|\, \lambda(u,t) \geq \frac{63}{64}\right\}, \tag{49}$$

where $\lambda(u,t)$ is defined in Eq. (2). By the above definition of $s(u)$ and since $u \in K_r$, we have $s(u) \geq 2r$. To simplify notation when the dependency on u is clear, we simply write $s := s(u)$. Let η denote the fraction of the surface of the ball $\mathbb{B}^n(u, s/2)$ that is not in K. Then using the fact that K is convex, we have

$$\text{vol}\left(\mathbb{B}^n(u, s) \setminus K\right) \geq \eta \cdot \text{vol}(\mathbb{B}^n(0, s)) - \text{vol}\left(\mathbb{B}^n(0, s/2)\right).$$

On the other hand, the definition of s implies that

$$\text{vol}\left(\mathbb{B}^n(u, s) \setminus K\right) \leq \frac{1}{64} \text{vol}(\mathbb{B}^n(0, s)).$$

We deduce that for $n \geq 9$,

$$\eta \leq \frac{1}{64} + 2^{-n} \leq \frac{3}{128}.$$

Take a line ℓ through u with direction uniformly distributed on the unit sphere. Then with probability at least $1 - 2\eta$, $\ell \cap \mathbb{B}^n(u, s/2) \subseteq K$.

Now, let p_u be the orthogonal projection of the point β on the line ℓ. Then $|p_u - u| = |u - \beta|\cos(\alpha)$ where α is the angle between ℓ and the line connecting u and β. An application of the spherical cap area upper bound as in Eq. (39) with $\cos(\alpha) = \frac{2\sqrt{2}}{\sqrt{n}}$ and a union bound implies that with probability at least $1 - 2\eta - \frac{1}{16}$, we have $\ell \cap \mathbb{B}^n(u, s/2) \subseteq K$ and $|p_u - u| \leq \sqrt{2} \cdot \frac{2\sqrt{2}}{\sqrt{n}} \leq \frac{4}{\sqrt{n}}$. Define $\tau := \min\left\{s, \frac{1}{8\sqrt{n}}\right\}$. Then

$$\mathbb{P}_{Y \sim P_{u \to \cdot}}\left(|Y - u| \leq \frac{\tau}{128} \,\Big|\, Y \in \ell\right) \overset{(i)}{\leq} \frac{\Gamma\left(\sqrt{n}\left(b + \frac{\tau}{256}\right)\right) - \Gamma\left(\sqrt{n}\left(b - \frac{\tau}{256}\right)\right)}{\Gamma\left(\sqrt{n}\left(b + \frac{s}{2}\right)\right) - \Gamma\left(\sqrt{n}\left(b - \frac{s}{2}\right)\right)}$$

$$\leq \frac{\Gamma\left(\sqrt{n}\left(b + \frac{\tau}{256}\right)\right) - \Gamma\left(\sqrt{n}\left(b - \frac{\tau}{256}\right)\right)}{\Gamma\left(\sqrt{n}\left(b + \frac{\tau}{2}\right)\right) - \Gamma\left(\sqrt{n}\left(b - \frac{\tau}{2}\right)\right)}$$

$$\leq \frac{\frac{\sqrt{n}\tau}{128}\gamma\left(\sqrt{n}\left(b - \frac{\tau}{256}\right)\right)}{\sqrt{n}\tau\gamma\left(\sqrt{n}\left(b + \frac{\tau}{2}\right)\right)}$$

$$\overset{(ii)}{\leq} \frac{1}{64},$$

where $b = |p_u - u|$, Γ is the cumulative density function of the standard Gaussian and γ is the density function of the standard Gaussian. (i) follows from looking at the one dimensional truncated Gaussian.

(ii) follows because $\sqrt{n}b \leq 4$, $\tau \leq \frac{1}{8\sqrt{n}}$, and a numerical calculation shows the ratio $\frac{\gamma(\sqrt{n}(b-\frac{\tau}{256}))}{\gamma(\sqrt{n}(b+\frac{\tau}{2}))} \leq 2$. For the unconditional probability, we have

$$\mathbb{P}_{Y \sim P_{u \to \cdot}}\left(|Y - u| \leq \frac{\tau}{128}\right) \leq (2\eta + \frac{1}{16}) \cdot 1 + (1 - 2\eta - \frac{1}{16}) \cdot \frac{1}{64} < \frac{1}{8}.$$

Hence,

$$F_u(\nu_{\beta,n}) \geq \frac{\tau}{128} \geq \frac{1}{128} \min\left\{2r, \frac{1}{8\sqrt{n}}\right\}.$$

$\square$

5.2.1 Proof of Lemma 8

Proof of Lemma 8. To simplify notation, let $\nu = \nu_{\beta,n}$. Consider the truncated density $\nu_\dagger$

$$\nu_\dagger(x) := \mathbf{1}_{K_r}(x) e^{-\frac{n}{2}|x-\beta|^2} \frac{1}{\nu(K_r)}$$

Note that since K_r is convex, $\nu_\dagger$ is still n-strongly-logconcave. Consequently, it satisfies the isoperimetric inequality (see Theorem 5.4 in [CV18]), for U_1, U_2, U_3 a partition of K,

$$\nu_\dagger(U_3) \geq \log 2 \cdot \sqrt{n} \cdot d(U_1, U_2) \cdot \nu_\dagger(U_1)\nu_\dagger(U_2).$$
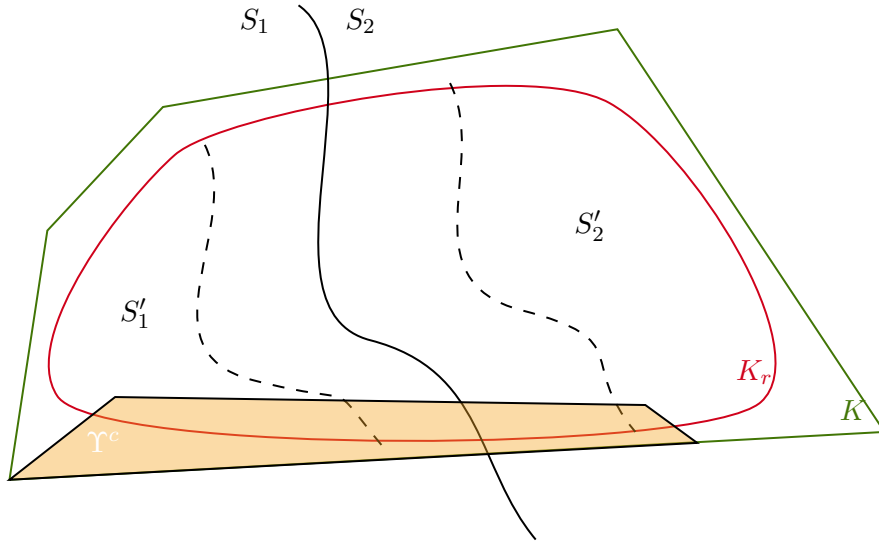

Figure 2: Illustration of the partition of K in conductance lower bound

Let $\varrho = \frac{r\sqrt{n}}{64C}$ where C is the constant from Lemma 23. Define the sets

$$S_1' := \left\{u \in S_1 \cap K_r \mid P_{u \to S_2}(\nu) < \frac{\varrho}{2}\right\}, \quad S_2' := \left\{u \in S_2 \cap K_r \mid P_{u \to S_1}(\nu) < \frac{\varrho}{2}\right\}.$$

There are two cases:

- Case 1: $\nu(S_1') \leq \nu(S_1 \cap K_r)/2$ or $\nu(S_2') \leq \nu(S_2 \cap K_r)/2$.

- Case 2: $\nu(S_i') \geq \nu(S_i \cap K_r)/2$ for $i = 1, 2$.

Case 1: Assuming that $\nu(S_1 \cap K_r \setminus S_1') \geq \nu(S_1 \cap K_r)/2$, we obtain

$$
\begin{aligned}
\int_{S_1} P_{x \to S_2}(\nu) d\nu(x) &\geq \int_{S_1 \cap K_r \setminus S_1'} P_{x \to S_2}(\nu) d\nu(x) \\
&\overset{(i)}{\geq} \frac{\varrho}{2} \nu(S_1 \cap K_r \setminus S_1') \\
&\geq \frac{\varrho}{4} \nu(S_1 \cap K_r).
\end{aligned}
$$

(i) follows from the definition of S_1'. The proof is similar for the case where the roles of S_1 and S_2 are switched, and using the reversibility of the kernel P which implies that $\int_{S_1} P_{x \to S_2}(\nu) d\nu(x) = \int_{S_2} P_{x \to S_1}(\nu) d\nu(x)$.

Case 2: For any $u \in S_1' \cap \Upsilon$ and $v \in S_2' \cap \Upsilon$, we have

$$
d_{\text{TV}} \left(P_{u \to \cdot}(\nu), P_{v \to \cdot}(\nu) \right) \geq P_{u \to S_1}(\nu) - P_{v \to S_1}(\nu) = 1 - P_{u \to S_2}(\nu) - P_{v \to S_1}(\nu) > 1 - \varrho.
$$

Lemma 24 implies that for $u \in S_1' \cap \Upsilon$ and $r \leq \frac{1}{16\sqrt{n}}$, we have

$$
F_u(\nu_{\beta,n}) \geq \frac{r}{64}. \tag{50}
$$

Together with Lemma 23, we have that $|u - v| \geq \Delta := \frac{r}{32\sqrt{n}}$. Since it holds for any pair of $u \in S_1' \cap \Upsilon$ and $v \in S_2' \cap \Upsilon$, its implies that

$$
d \left(S_1' \cap \Upsilon, S_2' \cap \Upsilon \right) \geq \Delta.
$$

We have

$$
\begin{aligned}
\int_{S_1} P_{x \to S_2}(\nu) d\nu(x) &= \frac{1}{2} \left(\int_{S_1} P_{x \to S_2} d\nu(x) + \int_{S_2} P_{x \to S_1} d\nu(x) \right) \\
&\geq \frac{1}{2} \left(\int_{S_1 \cap K_r \setminus S_1'} P_{x \to S_2} d\nu(x) + \int_{S_2 \cap K_r \setminus S_2'} P_{x \to S_1} d\nu(x) \right) \\
&\overset{(i)}{\geq} \frac{\varrho}{4} \left[\nu(S_1 \cap K_r \setminus S_1') + \nu(S_2 \cap K_r \setminus S_2') \right] \\
&= \frac{\varrho}{4} \nu(K_r \setminus (S_1' \cup S_2')).
\end{aligned}
$$

(i) follows from the definition of S_1' and S_2'. Note that the three sets $S_1' \cap \Upsilon$, $S_2' \cap \Upsilon$ and $K_r \setminus ((S_1' \cup S_2') \cap \Upsilon)$ form a partition of K_r. We have

$$
\begin{aligned}
\nu(K_r \setminus (S_1' \cup S_2')) + \delta &\geq \nu(K_r \setminus ((S_1' \cup S_2') \cap \Upsilon)) \\
&\overset{(i)}{\geq} \frac{\log(2) \cdot \sqrt{n}}{\nu(K_r)} \cdot d(S_1' \cap \Upsilon, S_2' \cap \Upsilon) \cdot \nu(S_1' \cap \Upsilon) \nu(S_2' \cap \Upsilon) \\
&\geq \frac{1}{2} \Delta \sqrt{n} \cdot \left(\nu(S_1' \cap \Upsilon) \right) \nu(S_2' \cap \Upsilon) \\
&\geq \frac{1}{2} \Delta \sqrt{n} \cdot \left(\nu(S_1') - \delta \right) \nu(S_2' \cap \Upsilon) \\
&\geq \frac{1}{2} \Delta \sqrt{n} \cdot \nu(S_1') \nu(S_2' \cap \Upsilon) - \frac{1}{2} \Delta \sqrt{n} \delta \\
&\geq \frac{1}{2} \Delta \sqrt{n} \cdot \nu(S_1') \nu(S_2') - \Delta \sqrt{n} \delta \\
&\overset{(ii)}{\geq} \frac{1}{8} \Delta \sqrt{n} \cdot \nu(S_1 \cap K_r) \cdot \nu(S_2 \cap K_r) - \Delta \sqrt{n} \delta.
\end{aligned}
$$

(i) applies the isoperimetric inequality for $\nu_\dagger$. (ii) applies the condition of Case 2. Combine the above two displays, we conclude there exists a universal constant $C' > 0$ such that

$$\int_{S_1} P_{u \to S_2}(\nu) d\nu(x) \geq \frac{r^2 \sqrt{n}}{C'} \left[\nu(S_1 \cap K_r) \cdot \nu(S_2 \cap K_r) - 8 \left(1 + \frac{32}{r} \right) \delta \right].$$

$\square$

References

[AC93]   James H Albert and Siddhartha Chib. Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association*, 88(422):669–679, 1993.

[AK91]   David 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, 1991.

[Ant07]   Irina A Antipova. Inversion of many-dimensional Mellin transforms and solutions of algebraic equations. *Sbornik: Mathematics*, 198(4):447, 2007.

[B$^+$97]   Keith Ball et al. An elementary introduction to modern convex geometry. *Flavors of geometry*, 31(1-58):26, 1997.

[BL02]   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. In *Inequalities*, pages 441–464. Springer, 2002.

[CDWY18] Yuansi Chen, Raaz Dwivedi, Martin J Wainwright, and Bin Yu. Fast MCMC sampling algorithms on polytopes. *The Journal of Machine Learning Research*, 19(1):2146–2231, 2018.

[CE22]   Yuansi Chen and Ronen Eldan. Localization schemes: A framework for proving mixing bounds for Markov chains. *arXiv preprint arXiv:2203.04163*, 2022.

[Che21]   Yuansi Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture. *Geometric and Functional Analysis*, 31(1):34–61, 2021.

[Cou17]   Benjamin Cousins. *Efficient high-dimensional sampling and integration*. PhD thesis, Georgia Institute of Technology, 2017.

[CV18]   Ben Cousins and Santosh Vempala. Gaussian cooling and $O^*(n^3)$ algorithms for volume and Gaussian volume. *SIAM Journal on Computing*, 47(3):1237–1273, 2018.

[DFK91]   Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *Journal of the ACM (JACM)*, 38(1):1–17, 1991.

[EAM22]   Ahmed El Alaoui and Andrea Montanari. An information-theoretic view of stochastic localization. *IEEE Transactions on Information Theory*, 2022.

[EKZ21]   Ronen Eldan, Frederic Koehler, and Ofer Zeitouni. A spectral condition for spectral gap: fast mixing in high-temperature Ising models. *Probability Theory and Related Fields*, pages 1–17, 2021.

[Eld13]   Ronen Eldan. Thin shell implies spectral gap up to polylog via a stochastic localization scheme. *Geometric and Functional Analysis*, 23(2):532–569, 2013.

[Eld22]   Ronen Eldan. Analysis of high-dimensional distributions using pathwise methods. In *Proceedings of ICM*, 2022.

[HH06]   Leonhard Held and Chris C Holmes. Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian analysis*, 1(1):145–168, 2006.

[JLLV21]   He Jia, Aditi Laddha, Yin Tat Lee, and Santosh Vempala. Reducing isotropy and volume to KLS: an $O(n^3\psi^2)$ volume algorithm. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, pages 961–974, 2021.

[JLV22]   Arun Jambulapati, Yin Tat Lee, and Santosh S Vempala. A slightly improved bound for the KLS constant. *arXiv preprint arXiv:2208.11644*, 2022.

[KL22]   Bo'az Klartag and Joseph Lehec. Bourgain's slicing problem and KLS isoperimetry up to polylog. *arXiv preprint arXiv:2203.15551*, 2022.

[KLS95]   Ravi Kannan, László Lovász, and Miklós Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13(3):541–559, 1995.

[KLS97]   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.

[KN09]   Ravi Kannan and Hariharan Narayanan. Random walks on polytopes and an affine interior point method for linear programming. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 561–570, 2009.

[LLV20]   Aditi Laddha, Yin Tat Lee, and Santosh Vempala. Strong self-concordance and sampling. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 1212–1222, 2020.

[LM00]   Beatrice Laurent and Pascal Massart. Adaptive estimation of a quadratic functional by model selection. *Annals of Statistics*, pages 1302–1338, 2000.

[Lov99]   László Lovász. Hit-and-run mixes fast. *Mathematical programming*, 86(3):443–461, 1999.

[LS90]   László Lovász and Miklós Simonovits. The mixing rate of Markov chains, an isoperimetric inequality, and computing the volume. In *Proceedings [1990] 31st annual symposium on foundations of computer science*, pages 346–354. IEEE, 1990.

[LS93]   László Lovász and Miklós Simonovits. Random walks in a convex body and an improved volume algorithm. *Random structures & algorithms*, 4(4):359–412, 1993.

[LV06a]   László Lovász and Santosh Vempala. Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 57–68. IEEE, 2006.

[LV06b]   László Lovász and Santosh Vempala. Hit-and-run from a corner. *SIAM Journal on Computing*, 35(4):985–1005, 2006.

[LV06c]   László Lovász and Santosh Vempala. Simulated annealing in convex bodies and an $O^*(n^4)$ volume algorithm. *Journal of Computer and System Sciences*, 72(2):392–417, 2006.

[LV07]   László Lovász and Santosh Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures & Algorithms*, 30(3):307–358, 2007.

[LV17]   Yin Tat Lee and Santosh S Vempala. Geodesic walks in polytopes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on theory of Computing*, pages 927–940, 2017.

[LV18a]   Yin Tat Lee and Santosh S Vempala. Convergence rate of Riemannian Hamiltonian Monte Carlo and faster polytope volume computation. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1115–1121, 2018.

[LV18b]   Yin Tat Lee and Santosh S Vempala. The Kannan-Lov\'asz-Simonovits conjecture. *arXiv preprint arXiv:1807.03465*, 2018.

[Mel96]   Hjalmar Mellin. *Über die fundamentale Wichtigkeit des Satzes von Cauchy für die Theorien der Gamma-und der hypergeometrischen Functionen*, volume 21. Societatis litterariae fennicae, 1896.

[MV19]   Oren Mangoubi and Nisheeth K Vishnoi. Faster polytope rounding, sampling, and volume computation via a sub-linear ball walk. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1338–1357. IEEE, 2019.

[Smi84]   Robert L Smith. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32(6):1296–1308, 1984.

[SW14]   Adrien Saumard and Jon A Wellner. Log-concavity and strong log-concavity: a review. *Statistics surveys*, 8:45, 2014.

[Tko12]   Tomasz Tkocz. An upper bound for spherical caps. *The American Mathematical Monthly*, 119(7):606–607, 2012.

[Tur71]   Valentin F Turchin. On the computation of multidimensional integrals by the Monte-Carlo method. *Theory of Probability & Its Applications*, 16(4):720–724, 1971.

[Vem05]   Santosh Vempala. Geometric random walks: a survey. *Combinatorial and computational geometry*, 52(573-612):2, 2005.

A Summary of properties of a logconcave distribution

Here is a list of well-known properties of logconcave distributions.

- The log-concavity of a measure is preserved by affine transformations and by marginalization, see Proposition 3.1 and Theorem 3.3 in [SW14].

- The strong log-concavity of a measure is preserved by affine transformations, by convolution (Theorem 3.7 in [SW14]) and by marginalization (Theorem 3.8 in [SW14]).

- The isoperimetric constant of a 1-dimensional isotropic logconcave density is lower bounded by $\log(2)/2 \approx 0.34$ (see Theorem 4.3 in [Vem05]).

- The maximal value of a 1-dimensional isotropic logconcave density p is bounded by 1 (Lemma 5.5 (a) in [LV07]). It is restated in Lemma 25.

- For an isotropic logconcave density p, $p(0) \geq 1/8$ (Lemma 5.5 (b) in [LV07]). It is restated in Lemma 26.

- An isotropic logconcave density has an exponential tail (Lemma 5.17 in [LV07]). It is restated in Lemma 27.

- Let a, b be the δ and $(1-\delta)$-quantile of p, then $p(a) \geq \frac{1}{8e}\delta$ and $p(b) \geq \frac{1}{8e}\delta$ as shown in Lemma 28.

- $|p'(a)| < \frac{2}{\delta}$ as shown in Lemma 29.

Lemma 25 (Lemma 5.5 (a) in [LV07]). *Let p be an isotropic logconcave density on $\mathbb{R}$. Then for any $x \in \mathbb{R}$,*

$$p(x) \leq 1.$$

Lemma 26 (Lemma 5.5 (b) in [LV07]). *Let p be an isotropic logconcave density on $\mathbb{R}$, then $p(0) \geq \frac{1}{8}$.*

See Lemma 5.5 in [LV07] for a proof of the above two lemmas.

Lemma 27. *For any isotropic logconcave density p in $\mathbb{R}^n$, and any $t > 0$, we have*

$$\mathbb{P}_{X \sim p}(|X| \geq t\sqrt{n}) \leq e^{-t+1}.$$

See Lemma 5.17 in [LV07] for a proof.

Lemma 28. *Let p be an isotropic logconcave density on $\mathbb{R}$ and let z be the $(1-\delta)$-quantile (that is, $\int_{-\infty}^z p(x)dz = 1 - \delta$), with $0 < \delta \leq 1/e$, then $p(z) \geq \frac{1}{8e}\delta$. Similarly, let w be the δ-quantile, then $p(w) \geq \frac{1}{8e}\delta$. Additionally, for any $\tilde{z} \in [w, z]$, we have $p(\tilde{z}) \geq \frac{1}{8e}\delta$.*

Proof of Lemma 28. According to Lemma 5.4 in [LV07], we have

$$z \geq 0.$$

Based on the tail of the logconcave density in Lemma 5.7 in [LV07], if $z \geq \log(e/\delta)$, then $z > 1$ and

$$\delta < e^{-z+1} \leq \delta,$$

which leads to a contradiction. Hence, $z < \log(e/\delta)$.

Suppose $p(z) < \frac{p(0)}{e}\delta < p(0)$. Applying the mean value theorem, there exists $y \in [0, z]$, such that

$$(\log p)'(y) = \frac{\log p(z) - \log p(0)}{z} \leq -1.$$

The derivative is non-increasing, as p is logconcave. Thus, for $x \geq z$, we have

$$(\log p)'(x) \leq -1.$$

Integrating from z to x, we obtain

$$p(x) \leq p(z)e^{-(x-z)}.$$

Integrating again from z to ∞, we obtain

$$\delta = \int_z^\infty p(x)dx \leq p(z) < \frac{p(0)}{e}\delta \leq \frac{1}{e}\delta,$$

which is a contradiction according to Lemma 25. Hence,

$$p(z) \geq \frac{p(0)}{e}\delta \geq \frac{1}{8e}\delta,$$

where the last step follows from Lemma 26.

The proof of the δ-quantile is similar. Additionally, since $(\log p)'(z) \leq -1$, $(\log p)'(w) \geq 1$ and the derivative is non-increasing, we conclude that for any $\tilde{z} \in [w, z]$, we have $p(\tilde{z}) \geq \frac{1}{8e}\delta$. $\qquad\square$

Lemma 29. *Let p be an isotropic logconcave density on $\mathbb{R}$ and let z be the $(1 - \delta)$-quantile. That is, $\int_{-\infty}^z p(x)dz = 1 - \delta$. Suppose $0 < \delta \leq 1/e$, then there exists a universal constant $c > 0$ such that $p'(z) \geq -\frac{2}{\delta}$. By symmetry, a similar result holds for the δ-quantile.*

Proof of Lemma 29. Suppose $p'(z) < -\frac{2}{\delta}$. Then because $0 \leq p(z) \leq 1$,

$$(\log p)'(z) = \frac{p'(z)}{p(z)} < -\frac{2}{\delta}.$$

The derivative is non-increasing as p is logconcave. Thus, for $x \geq z$, we have

$$(\log p)'(x) < -\frac{2}{\delta}.$$

Integrating twice, we obtain

$$\delta = \int_z^\infty p(x)dx \leq \int_z^\infty e^{-\frac{2}{\delta}(x-z)}dx \leq \frac{\delta}{2},$$

which is a contradiction.

$\qquad\square$