Stochastic Localization and Lee and Vempala's result on the KLS conjecture

Yunwei Ren*

May 16, 2024

Abstract

The purpose of this note is to provide a short introduction to the stochastic localization technique introduced by [Eld13] and the progress made by [LV19] on the KLS conjecture [KLS95], where this technique plays a central role.

1 Introduction

The Kannan-Lovász-Simonovits (KLS) conjecture says that the Cheeger constant of any isotropic log-concave distribution on $\mathbb{R}^d$ can be lower bounded by some universal constant. Formally, for a log-concave distribution $\nu \in \mathscr{M}_1(\mathbb{R}^d)$ and measurable set $A \subset \mathbb{R}^d$, we define the surface measure of A as

$$\nu^+(\partial A) := \limsup_{\varepsilon \to 0+} \frac{\nu(A_\varepsilon \setminus A)}{\varepsilon},$$

where A_ε is the ε-neighborhood of A. The Cheeger constant is defined as the minimum of the ratio between the surface measure and the volume, that is,

$$\psi_\nu := \inf_{A \subset \mathbb{R}^d} \frac{\nu^+(\partial A)}{\nu(A)(1 - \nu(A))}. \tag{1}$$

The KLS conjecture can be formally stated as follows.

Conjecture 1.1 (KLS conjecture, [KLS95])**.** *There exists a universal constant $c > 0$ such that for any log-concave distribution $\nu \in \mathscr{M}_1(\mathbb{R}^d)$ with an identity covariance matrix, we have $\psi_\nu \geq c$.*

*Princeton University, `yunwei.ren@princeton.edu`

This conjecture unifies and/or implies many conjectures and results in different areas of mathematics and computer science including convex geometry, probability and sampling algorithms. One may refer to [LV18] for a survey on the connection of KLS with other conjectures and results.

The original bound given by [KLS95] is $\psi_\nu \gtrsim d^{-1/2}$. Recently, this bound has been greatly improved. In [LV19], the author show that $\psi_\nu \gtrsim d^{-1/4}$, [Che21] improves this bound to $d^{-o(1)}$, and later on this bound is further improved to poly log d by [KL22]. The central technique behind these progresses is the stochastic localization, first introduced by Eldan in [Eld13].

In this note, we will first define stochastic localization and discuss its basic properties and the general strategy of applying it to the KLS conjecture (cf. Section 2). After that, we will reproduce the $d^{-1/4}$ bound given by [LV19] in Section 3.

2 Stochastic localization and the reduction

In this section, we define the stochastic localization process and discuss the basic properties of it. The construction we present here comes from [EAM22]. After that, we show that in order to lower bound ψ_ν, it suffices to upper bound certain covariance matrices along the stochastic localization process.

2.1 Construction and properties of the stochastic localization

One basic strategy in probability theory is decomposing a distribution $\nu \in \mathscr{M}_1(\mathbb{R}^d)$ into simpler components and analyze each of these components. Examples include tensorization and pinning ([Mon08]). The way stochastic localization decomposes the distribution is constructing a probability measure-valued process $(\nu_t)_t$ such that $(\nu_t(A))_t$ is a martingale for every measurable $A \subset \mathbb{R}^d$, and then we can write $\nu = \mathbb{E}\, \nu_t$ where the expectation is taken over the randomness of $(\nu_t)_t$.

Formally, let $X \sim \nu$ and $(W_t)_t$ be an independent Brownian motion. For $t \geq 0$, we set

$$Y_t = tX + W_t,$$

and let ν_t be the conditional distribution of X given the value of Y_t. That is,

$$\nu_t(\cdot) := \mathbb{P}\left[X \in \cdot \mid Y_t\right].$$

Clear that $(\nu_t)_t$ is a martingale with $\nu_0 = \nu$. Moreover, note that $Y_t | X \sim \mathcal{N}(tX, t)$. Hence, by Bayes' rule, we have

$$\nu_t(\mathrm{d}x) \propto \exp\left(-\frac{1}{2t}\|Y_t - tx\|_2^2\right)\nu(x) \propto \exp\left(\langle Y_t, x\rangle - \frac{t}{2}\|x\|^2\right)\nu(x). \tag{2}$$

In particular, this equation implies that when ν is log-concave, ν_t will be t-strongly log-concave. This observation is important when applying stochastic localization to the KLS conjecture. Now, we write down the SDE that controls the evolution of ν_t.

Lemma 2.1 (Theorem 2 of [EAM22])**.** *Let* $F_t(x) := \nu_t(\mathrm{d}x)/\nu(\mathrm{d}x)$. *Then, there exists a standard Brownian motion* B *such that, for every* $x \in \mathbb{R}^d$,

$$F_0(x) = 1, \quad \mathrm{d}F_t(x) = F_t(x) \langle x - a_t, \mathrm{d}B_t \rangle \,, \tag{3}$$

where $a_t := \int x \nu_t(\mathrm{d}x)$ *is the center of mass of* ν_t.

Remark. (3) is how stochastic localization was originally defined. Intuitively, it means at each step, we tilt the distribution via a random linear function. ♣

Proof. Write $F_t(\mathrm{d}x) = \frac{1}{Z_t} \exp\left(\langle Y_t, x\rangle - \frac{t}{2}\|x\|^2\right)$, where Z_t is the normalizing constant. We have

$$\mathrm{d}\log F_t(x) = \langle \mathrm{d}Y_t, x\rangle - \frac{1}{2}\|x\|^2 \mathrm{d}t - \mathrm{d}\log Z_t.$$

First, consider the normalizing constant. We have

$$\begin{aligned}
\mathrm{d}Z_t &= \mathrm{d}\int \exp\left(\langle Y_t, x\rangle - \frac{t}{2}\|x\|^2\right)\nu(\mathrm{d}x) \\
&= \int \exp\left(\langle Y_t, x\rangle - \frac{t}{2}\|x\|^2\right)\left(\langle \mathrm{d}Y_t, x\rangle - \frac{\mathrm{d}t}{2}\|x\|^2 + \frac{1}{2}\mathrm{d}[\langle Y_\cdot, x\rangle]_t\right)\nu(\mathrm{d}x) \\
&= \left\langle \mathrm{d}Y_t, \int x \exp\left(\langle Y_t, x\rangle - \frac{t}{2}\|x\|^2\right)\nu(\mathrm{d}x)\right\rangle \\
&= Z_t \langle \mathrm{d}Y_t, a_t\rangle \,.
\end{aligned}$$

Hence,

$$\mathrm{d}\log Z_t = \frac{\mathrm{d}Z_t}{Z_t} - \frac{1}{2}\frac{\mathrm{d}[Z]_t}{Z_t^2} = \langle \mathrm{d}Y_t, a_t\rangle - \frac{1}{2}\|a_t\|^2 \mathrm{d}t.$$

Thus, for $\log F_t(x)$, we have

$$\begin{aligned}
\mathrm{d}\log F_t(x) &= \langle \mathrm{d}Y_t, x\rangle - \frac{1}{2}\|x\|^2 \mathrm{d}t - \langle \mathrm{d}Y_t, a_t\rangle + \frac{1}{2}\|a_t\|^2 \mathrm{d}t \\
&= \langle \mathrm{d}Y_t - a_t \mathrm{d}t, x - a_t\rangle - \frac{1}{2}\|x - a_t\|^2 \,.
\end{aligned}$$

Then, we compute

$$\begin{aligned}
\mathrm{d}F_t(x) = \mathrm{d}\exp(\log F_t(x)) &= F_t(x)\left(\mathrm{d}\log F_t(x) + \frac{1}{2}\mathrm{d}[\log F_\cdot(x)]_t\right) \\
&= F_t(x)\left(\mathrm{d}\log F_t(x) + \frac{1}{2}\|x - a_t\|^2\right) \\
&= F_t(x)\langle x - a_t, \mathrm{d}Y_t - a_t \mathrm{d}t\rangle \,.
\end{aligned}$$

To complete the proof, put $M_t = Y_t - \int_0^t a_u \mathrm{d}u$ and note that for any $s \le t$,

$$\begin{aligned}
\mathbb{E}[M_t \mid \mathcal{F}_s] &= M_s + \mathbb{E}\left[Y_t - Y_s - \int_s^t \mathbb{E}[X|\mathcal{F}_u]\mathrm{d}u \,\middle|\, \mathcal{F}_s\right] \\
&= M_s + \mathbb{E}[Y_t \mid \mathcal{F}_s] - Y_s - \int_s^t \mathbb{E}[X|\mathcal{F}_s]\mathrm{d}u \\
&= M_s + \mathbb{E}[sX + W_t + (t-s)X \mid \mathcal{F}_s] - Y_s - (t-s)a_s = M_s,
\end{aligned}$$

Therefore, M is a martingale with quadratic variation $[M]_t = t$. Thus, by Lévy's characterization, it is a Brownian motion. □

2.2 Reducing the problem to estimating the covariance matrix

In this subsection, we show that in order to lower bound the Cheeger constant, it suffices to upper bound $\mathbb{E}\int_0^t \|\mathrm{Cov}(\nu_s)\|_{\mathrm{OP}}\,\mathrm{d}s$. The presentation here mostly follows Section 4.2 of [Eld23].

Recall that (2) implies that if ν is log-concave, then ν_t is a t-strongly log-concave distribution, which is known to have a dimension-independent Cheeger constant (cf. Theorem 2.2). Hence, we can apply Theorem 2.2 to ν_t and bound the difference between ν and ν_t using the SDE formulation (3) of $(\nu_t)_t$.

Theorem 2.2 ([BL96])**.** *If $\mu \in \mathscr{M}_1(\mathbb{R}^d)$ is α-strongly log-concave for some $\alpha > 0$, then, for any measurable $A \subset \mathbb{R}^d$, we have*

$$\mu^+(\partial A) \ge \sqrt{\alpha}\mu(A)(1-\mu(A)).$$

Let $M_t := \nu_t(A)$. Using the martingale property of $(\nu_t)_t$, we can write

$$\begin{aligned}
\nu^+(\partial A) = \mathbb{E}\,\nu_t^+(\partial A) &\ge \sqrt{t}\,\mathbb{E}[\nu_t(A)(1-\nu_t(A))] \\
&= \sqrt{t}\left(\nu(A)(1-\nu(A)) - \mathbf{Var}\,M_t\right).
\end{aligned} \tag{4}$$

As a result, it suffices to upper bound $\mathbf{Var}\,M_t$, which is equal to $\mathbb{E}[M]_t$, where $([M]_t)_t$ denotes the quadratic variation of M. This can be done using (3). Formally, we have the following lemma.

Lemma 2.3. *Let $\nu \in \mathscr{M}_1(\mathbb{R}^d)$ be a log-concave distribution and $(\nu_t)_t$ the associated stochastic localization process. Suppose that for some $t > 0$ and $\alpha \in (0, 1/8]$, we have*

$$\mathbb{E}\int_0^t \|\mathrm{Cov}(\nu_s)\|_{\mathrm{OP}}\,\mathrm{d}s \le \alpha. \tag{5}$$

Then, we have, for any A with $\nu(A)(1-\nu(A)) \ge 2\alpha$, we have

$$\nu^+(\partial A) \ge \frac{1}{2}\sqrt{t}\nu(A)(1-\nu(A)).$$

Moreover, under mild regularity conditions on ν, this implies $\psi_\nu \ge \sqrt{t}/2$.

Remark. With this lemma, our goal becomes choosing a t as large as possible, without violating (5). ♣

Proof. By (3), we have $\mathrm{d}M_t = \int_A \langle x - a_t, \mathrm{d}B_t\rangle \, \nu_t(\mathrm{d}x)$. Therefore,

$$\begin{aligned} \mathrm{d}[M]_t = \left\| \int_A (x-a_t)\nu_t(\mathrm{d}x) \right\|_2^2 \mathrm{d}t \le \int \|x-a_t\|_2^2 \, \nu_t(\mathrm{d}x)\mathrm{d}t \\ &= \sup_{\theta \in \mathbb{S}^{d-1}} \int \langle x-a_t, \theta\rangle^2 \, \nu_t(\mathrm{d}x)\mathrm{d}t \\ &= \sup_{\theta \in \mathbb{S}^{d-1}} \theta^\top \left(\int (x-a_t)(x-a_t)^\top \nu_t(\mathrm{d}x) \right) \theta \mathrm{d}t \\ &= \|\mathrm{Cov}(\nu_t)\|_{\mathrm{OP}} \, \mathrm{d}t. \end{aligned}$$

Integrate both sides from 0 to t, take expectation, and we obtain

$$\mathbb{E}[M]_t \le \mathbb{E} \int_0^t \|\mathrm{Cov}(\nu_s)\|_{\mathrm{OP}} \, \mathrm{d}s \le \alpha.$$

Combining these with (4), we get

$$\nu^+(\partial A) \ge \sqrt{t}\,(\nu(A)(1-\nu(A)) - \alpha) \ge \frac{1}{2}\sqrt{t}\nu(A)(1-\nu(A)),$$

for any A with $\nu(A)(1-\nu(A)) \ge 2\alpha \ge 1/4$. By Theorem 25 of [LV19] (or Theorem 1.8 of [Mil09]), under mild regularity conditions on ν, this implies $\psi_\nu \ge \sqrt{t}/2$. □

In [LV19], a variant of the above lemma is given, in which (5) is replaced by a probability bound. This version allows us to focus on part of the probability space and ignore certain potential bad cases.

Lemma 2.4. *Under mild regularity conditions on ν, if for some $t > 0$, we have*

$$\mathbb{P}\left[\int_0^t \|\mathrm{Cov}(\nu_s)\|_{\mathrm{OP}} \, \mathrm{d}s \le \frac{1}{64} \right] \ge \frac{3}{4}, \tag{6}$$

then we have $\psi_\nu \gtrsim \sqrt{t}$.

Proof. Following the calculation in (4), we have

$$\nu^+(\partial A) \ge \sqrt{t}\,\mathbb{E}\left[\nu_t(A)(1-\nu_t(A))\right] \ge \frac{\sqrt{t}}{16}\,\mathbb{P}\left[\frac{1}{4} \le \nu_t(A) \le \frac{3}{4}\right]. \tag{7}$$

By Theorem 1.8 of [Mil09], it suffices to consider the case where $M_0 = \nu(A) = 1/2$. By the Dambis-Dubins-Schwarz theorem, there exists a Brownian motion $\tilde{B}$ such that

$M_t - M_0 = \tilde{B}_{[M]_t}$. Meanwhile, for the quadratic variation of M, by the proof of Lemma 2.3, we have $\mathrm{d}[M]_t \le \|\mathrm{Cov}(\nu_t)\|_{\mathrm{OP}}\,\mathrm{d}t$. Thus,

$$\begin{aligned}
\mathbb{P}\left[\frac{1}{4} \le \nu_t(A) \le \frac{3}{4}\right] &= \mathbb{P}\left[\left|\tilde{B}_{[M]_t}\right| \le \frac{1}{4}\right] \\
&\ge 1 - \mathbb{P}\left[\max_{s\le 1/64} |\tilde{B}_s| \ge \frac{1}{4}\right] - \mathbb{P}\left[[M]_t \ge \frac{1}{64}\right] \\
&\ge 1 - 4\,\mathbb{P}\left[\tilde{B}_{1/64} \ge \frac{1}{4}\right] - \mathbb{P}\left[[M]_t \ge \frac{1}{64}\right] \\
&\ge \frac{9}{10} - \mathbb{P}\left[[M]_t \ge \frac{1}{64}\right],
\end{aligned}$$

where the second inequality comes from the reflection principle of Brownian motion and the last from the concentration of Gaussian distribution. Combine this with (7) and condition (6), and we get

$$\begin{aligned}
\nu^+(\partial A) &\ge \frac{\sqrt{t}}{16}\left(\frac{9}{10} - \mathbb{P}\left[[M]_t \ge \frac{1}{64}\right]\right) \\
&\ge \frac{\sqrt{t}}{16}\left(\frac{9}{10} - \mathbb{P}\left[\int_0^t \|\mathrm{Cov}\,\nu_s\|_{\mathrm{OP}}\,\mathrm{d}s \ge \frac{1}{64}\right]\right) \gtrsim \sqrt{t}.
\end{aligned}$$

□

3 Lee and Vempala's bound on the covariance matrix

For notational simplicity, let $A_t := \mathrm{Cov}(\nu_t)$. Instead of controlling $\|A_t\|_{\mathrm{OP}}$, we consider the $\mathrm{Tr}(A_t^2)$. This function, unlike $\|A_t\|_{\mathrm{OP}}^2$, is smooth in A_t, and we have

$$\int_0^t \|A_s\|_{\mathrm{OP}}\,\mathrm{d}s \le \int_0^t \sqrt{\mathrm{Tr}(A_s^2)}\mathrm{d}s \le t\sqrt{\max_{s\in[0,t]} \mathrm{Tr}(A_s^2)}.$$

For the dynamics of A_t and $\mathrm{Tr}(A_t^2)$, we have the following lemmas, both of which follow from standard stochastic calculus and the proofs are deferred to the end of this section.

Lemma 3.1 (Dynamics of A_t)**.** *We have*

$$\mathrm{d}A_t = \int (x-a_t)(x-a_t)^\top \langle x - a_t, \mathrm{d}B_t\rangle\, \nu_t(\mathrm{d}x) - A_t^2\mathrm{d}t.$$

Lemma 3.2 (Dynamics of $\mathrm{Tr}(A_t^2)$)**.** *We have*

$$\begin{aligned}
\mathrm{d}\,\mathrm{Tr}\,A_t^2 = 2&\left[\int \langle A_t(x-a_t), x-a_t\rangle\,(x-a_t)\nu_t(\mathrm{d}x)\right]^\top \mathrm{d}B_t \\
&- 2\,\mathrm{Tr}(A_t^3)\mathrm{d}t + \iint \langle x-a_t, y-a_t\rangle^3\, \nu_t(\mathrm{d}x)\nu_t(\mathrm{d}y)\mathrm{d}t
\end{aligned}$$

To proceed, we upper bound the RHS using $\operatorname{Tr} A_t^2$ and apply Gronwall's lemma. Since $A_t \succeq 0$, it suffices to consider the first and third terms on the RHS[1]. For the last term, we have the following general bound.

Lemma 3.3. *For any log-concave distribution μ with mean a and covariance matrix A, we have*

$$\iint |\langle x-a, y-a\rangle|^3 \,\mu(\mathrm{d}x)\mu(\mathrm{d}y) \lesssim \operatorname{Tr}(A^2)^{3/2}.$$

Proof. Assume w.l.o.g. that $a = 0$. For fix x and random $y \sim \mu$, $\langle x, y\rangle$ is also log-concave and therefore, by Theorem 5.22 of [LV07], we have

$$\int |\langle x, y\rangle|^3 \mu(\mathrm{d}y) \lesssim \left(\int \langle x, y\rangle^2 \,\mu(\mathrm{d}y)\right)^{3/2} = \left\|A^{1/2}x\right\|_2^3.$$

Therefore, we have

$$\begin{aligned}\iint |\langle x-a, y-a\rangle|^3 \,\mu(\mathrm{d}x)\mu(\mathrm{d}y) &\lesssim \int \left\|A^{1/2}x\right\|_2^3 \mu(\mathrm{d}x)\\ &\lesssim \left(\int \left\|A^{1/2}x\right\|_2^2 \mu(\mathrm{d}x)\right)^{3/2} = \operatorname{Tr}(A^2)^{3/2},\end{aligned}$$

where the second inequality comes from the Jensen's inequality. □

For the first term, we have the following bound on the coefficient.

Lemma 3.4. *Put $v_t := \int \langle A_t(x-a_t), x-a_t\rangle\,(x-a_t)\nu_t(\mathrm{d}x)$. We have $\|v_t\|_2 \lesssim \operatorname{Tr}(A_t^2)^{5/4}$.*

Proof. Let $A_t = \sum_{i=1}^d \lambda_i v_i v_i^\top$ be the spectral decomposition of A_t. For notational simplicity, also define $w_i = A_t^{1/2} v_i$ and $y = A_t^{-1/2}(x-a_t)$. Then, we can write

$$v_t = \sum_{i=1}^d \lambda_i \int \langle v_i, x-a_t\rangle^2\,(x-a_t)\nu_t(\mathrm{d}x) = \sum_{i=1}^d \lambda_i \mathop{\mathbb{E}}_{x\sim\nu_t}\left[\langle w_i, y\rangle^2\,(x-a_t)\right].$$

Then, we compute

$$\begin{aligned}\|v_t\| = \sup_{\xi\in\mathbb{S}^{d-1}} \langle v_t, \xi\rangle &= \sup_{\xi\in\mathbb{S}^{d-1}} \sum_{i=1}^d \lambda_i \mathop{\mathbb{E}}_{x\sim\nu_t}\left[\langle w_i, y\rangle^2 \langle \xi, x-a_t\rangle\right]\\ &\le \sup_{\xi\in\mathbb{S}^{d-1}} \sum_{i=1}^d \lambda_i \sqrt{\mathop{\mathbb{E}}_{x\sim\nu_t} \langle w_i, y\rangle^4}\sqrt{\mathop{\mathbb{E}}_{x\sim\nu_t} \langle \xi, x-a_t\rangle^2}.\end{aligned}$$

[1]One may attempt to take expectation on both sides to get rid of the first term and use Lemma 2.3. Unfortunately, this strategy will not work smoothly since the bound we will obtain for the third term is $\operatorname{Tr}(A^2)^{3/2}$. When we take expectation, it becomes $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\operatorname{Tr} A_t^2 \lesssim \mathbb{E}(\operatorname{Tr} A_t^2)^{3/2}$ and there is no easy way to move the 3/2 exponent out of the expectation.

For the second expectation, we have $\mathbb{E}_{x\sim\nu_t}\langle \xi, x-a_t\rangle^2 = \xi^\top A_t \xi_i \le \|A_t\|_{\mathrm{OP}}$. For the first expectation, note that $\langle w_i, y\rangle$ follows a log-concave distribution, whence again by Theorem 5.22 of [LV07], we have $\mathbb{E}_{x\sim\nu_t}\langle w_i, y\rangle^4 \lesssim \left(\mathbb{E}_{x\sim\nu_t}\langle w_i, y\rangle^2\right)^2$. Therefore, we have

$$\begin{aligned}\|v_t\| \lesssim \|A\|_{\mathrm{OP}}^{1/2}\sum_{i=1}^d \lambda_i \mathop{\mathbb{E}}_{x\sim\nu_t}\langle w_i, y\rangle^2 &= \|A\|_{\mathrm{OP}}^{1/2}\sum_{i=1}^d \lambda_i \mathop{\mathbb{E}}_{x\sim\nu_t}\langle v_i, x-a_t\rangle^2 \\ &= \|A\|_{\mathrm{OP}}^{1/2}\sum_{i=1}^d \lambda_i \operatorname{Tr}(A_t^{1/2} v_i v_i A_t^{1/2}) \\ &= \|A\|_{\mathrm{OP}}^{1/2}\operatorname{Tr}(A_t^2).\end{aligned}$$

To complete the proof, it suffices to note that $\|A\|_{\mathrm{OP}}^{1/2} \le \operatorname{Tr}(A_t^2)^{1/4}$. □

Theorem 3.5 (Theorem 12 of [LV19])**.** *For any isotropic log-concave distribution ν, $\psi_\nu \gtrsim n^{-1/4}$.*

Proof. For notational simplicity, put $\Phi_t = \operatorname{Tr}(A_t^2)$. By Lemma 3.2, Lemma 3.3 and Lemma 3.4, we have

$$\mathrm{d}\Phi_t \lesssim \Phi_t^{3/2}\mathrm{d}t + v_t^\top \mathrm{d}B_t$$

By Lemma 3.4, the third should approximately be "$\Phi_t^{5/4}|\mathrm{d}B_t| \approx \Phi_t^{5/4}\sqrt{t}$". This suggests the choice of $t \sim 1/\sqrt{\Phi_0} = 1/\sqrt{d}$. Now, to decouple these two terms, consider $f(\phi) = -(\phi+d)^{-\theta}$, where $\theta > 0$ is a parameter to be determined later. We have $f'(\phi) = \theta(\phi+d)^{-\theta-1}$ and $f''(\phi) = -\theta(\theta+1)(\phi+d)^{-\theta-2}$. Hence, by Itô's formula, we have

$$\begin{aligned}\mathrm{d}f(\Phi_t) &= \theta(\Phi_t+d)^{-\theta-1}\mathrm{d}\Phi_t - \frac{\theta(\theta+1)}{2}(\Phi_t+d)^{-\theta-2}\mathrm{d}[\Phi]_t \\ &\lesssim \theta\frac{\Phi_t^{3/2}}{(\Phi_t+d)^{\theta+1}}\mathrm{d}t + \theta\frac{v_t^\top \mathrm{d}B_t}{(\Phi_t+d)^{\theta+1}} - \frac{\theta(\theta+1)}{2}\frac{\|v_t\|^2}{(\Phi_t+d)^{\theta+2}}\mathrm{d}t \\ &\lesssim \theta\frac{\Phi_t^{3/2}}{(\Phi_t+d)^{\theta+1}}\mathrm{d}t - \frac{\theta(\theta+1)}{2}\frac{\Phi_t^{5/2}}{(\Phi_t+d)^{\theta+2}}\mathrm{d}t + \theta\frac{v_t^\top \mathrm{d}B_t}{(\Phi_t+d)^{\theta+1}}.\end{aligned}$$

Choose $\theta = 0.5$ so that the coefficient of the first two terms become constants. Then, we have

$$f(\Phi_t) - f(\Phi_0) \le Ct + \int_0^t \frac{v_s^\top \mathrm{d}B_s}{(\Phi_s+d)^{1.5}} =: Ct + N_t,$$

for some constant $C > 0$. Now, we bound the probability that the martingale N is large. The proof is similar to the one of Lemma 2.4. For the martingale N, we have

$$\mathrm{d}[N]_t = \frac{\|v_t\|^2}{(\Phi_t+d)^3} \lesssim \frac{\Phi_t^{2.5}}{(\Phi_t+d)^3} \le \frac{1}{\sqrt{d}}.$$

Also, by the Dambis-Dubins-Schwarz theorem, there exists some Brownian motion $\tilde{B}$ such that $N_t = \tilde{B}_{[N]_t}$. Therefore, for any $\gamma > 0$ and some universal constant $C_1 > 0$,

$$\mathbb{P}\left[\max_{s\le t} N_t \ge \gamma\right] \le \mathbb{P}\left(\max_{s\le C_1 t/\sqrt{d}} \tilde{B}_s \ge \gamma\right) \le 2\,\mathbb{P}\left(\tilde{B}_{C_1 t/\sqrt{d}} \ge \gamma\right) \le 2\exp\left(-\frac{\gamma^2\sqrt{d}}{2C_1 t}\right).$$

In other words, with probability at least $1 - 2\exp\left(-\frac{\gamma^2\sqrt{d}}{2C_1 t}\right)$, we have

$$f(\Phi_t) \le f(\Phi_0) + Ct + \gamma = -\frac{1}{\sqrt{2d}} + Ct + \gamma.$$

Choose $\gamma = \frac{0.25}{\sqrt{2d}}$. Then, note that we can choose $t = c/\sqrt{d}$ for some sufficiently small constant $c > 0$ such that

$$-\frac{1}{\sqrt{\Phi_s + d}} = f(\Phi_s) \le -\frac{1}{2}\frac{1}{\sqrt{2d}} \quad \text{for all } s \le t \text{ with probability at least } 0.99.$$

This implies that $\mathbb{P}[\max_{t\le c/\sqrt{d}} \mathrm{Tr}(A_t^2) \le 7d] \ge 0.99$. As a result, for some potential smaller $c > 0$, we have

$$\mathbb{P}\left[\int_0^{c/\sqrt{d}} \|A_s\|_{\mathrm{OP}}\,\mathrm{d}s \le \frac{1}{64}\right] \ge \mathbb{P}\left[\frac{c}{\sqrt{d}}\sqrt{\max_{s\in[0,c/\sqrt{d}]} \mathrm{Tr}(A_s^2)} \le \frac{1}{64}\right] \ge \frac{3}{4}.$$

Thus, by Lemma 2.4, we have $\psi_\nu \gtrsim d^{-1/4}$. □

Deferred proofs

Proof of Lemma 3.1. Recall that $A_t = \int (x-a_t)(x-a_t)^\top \nu_t(\mathrm{d}x)$. First, for a_t, we have

$$\begin{aligned}\mathrm{d}a_t = \int x\mathrm{d}F_t(x)\nu(\mathrm{d}x) &= \int x\,\langle x - a_t, \mathrm{d}B_t\rangle\,\nu_t(\mathrm{d}x)\\ &= \left(\int x(x-a_t)^\top \nu_t(\mathrm{d}x)\right)\mathrm{d}B_t = A_t\mathrm{d}B_t.\end{aligned}$$

Then, we compute

$$\begin{aligned}\mathrm{d}A_t &= \mathrm{d}\int (x-a_t)(x-a_t)^\top F_t(x)\nu(\mathrm{d}x)\\ &= -\mathrm{d}a_t\int (x-a_t)^\top F_t(x)\nu(\mathrm{d}x) - \int (x-a_t)F_t(x)\nu(\mathrm{d}x)(\mathrm{d}a_t)^\top\\ &\quad + \int (x-a_t)(x-a_t)^\top \mathrm{d}F_t(x)\nu(\mathrm{d}x) + \mathrm{d}[a_\cdot, a_\cdot]_t\int F_t(x)\nu(\mathrm{d}x)\\ &\quad - \int (x-a_t)\mathrm{d}[a_\cdot, F_\cdot(x)]_t^\top \nu(\mathrm{d}x) - \int \mathrm{d}[a_\cdot, F_\cdot(x)]_t(x-a_t)^\top \nu(\mathrm{d}x).\end{aligned}$$

Note that the first two terms are 0. For the second line, we have

$$\int (x - a_t)(x - a_t)^\top \mathrm{d}F_t(x)\nu(\mathrm{d}x) = \int (x - a_t)(x - a_t)^\top \langle x - a_t, \mathrm{d}B_t\rangle \nu_t(\mathrm{d}x),$$

and for each $i, j \in [d]$,

$$\mathrm{d}[a_i, a_j]_t = \left\langle A_i, A_j \right\rangle \mathrm{d}t = [A_t^2]_{i,j}\mathrm{d}t.$$

For the last two terms, note that

$$\int (x - a_t)\mathrm{d}[a_\cdot, F_\cdot(x)]_t^\top \nu(\mathrm{d}x) = \int (x - a_t)(x - a_t)^\top A_t \mathrm{d}t \nu_t(\mathrm{d}x) = A_t^2 \mathrm{d}t.$$

Combine these together and we complete the proof. □

Proof of Lemma 3.2. First, we compute the first and second order derivatives of $\Phi(A) = \mathrm{Tr}(A^2)$. Let Δ be a perturbation with the same shape of A. We have

$$\Phi(A + \Delta) = \mathrm{Tr}((A + \Delta)^2) = \mathrm{Tr}(A^2) + 2\,\mathrm{Tr}(A\Delta) + \mathrm{Tr}(\Delta^2).$$

Hence, we first and second differentials are $\Phi'(A)[\Delta] = 2\,\mathrm{Tr}(A\Delta)$ and $\Phi''(A)[\Delta_1, \Delta_2] = 2\,\mathrm{Tr}(\Delta_1\Delta_2)$. Then, by Itô's formula and Lemma 3.1, we have

$$\mathrm{d}\,\mathrm{Tr}(A_t^2) = 2\,\mathrm{Tr}(A_t \mathrm{d}A_t) + \frac{1}{2} \sum_{i,j\in[d]} \mathrm{d}[A_{i,j}, A_{j,i}]_t.$$

The first term is equal to

$$2\,\mathrm{Tr}(A_t \mathrm{d}A_t) = 2 \int \langle A_t(x - a_t), (x - a_t)\rangle \langle x - a_t, \mathrm{d}B_t\rangle \nu_t(\mathrm{d}x) - 2\,\mathrm{Tr}(A_t^3)\mathrm{d}t.$$

For the second term, note that

$$\mathrm{d}[A_t]_{i,j} = \int (x - a_t)_i (x - a_t)_j \langle x - a_t, \mathrm{d}B_t\rangle \nu_t(\mathrm{d}x) - [A_t^2]_{i,j}\mathrm{d}t,$$

and therefore,

$$\begin{aligned}
&\sum_{i,j} \mathrm{d}[A_{i,j}, A_{j,i}]_t \\
&= \sum_{i,j} \iint (x - a_t)_i (x - a_t)_j (y - a_t)_i (y - a_t)_j \langle x - a_t, y - a_t\rangle \nu_t(\mathrm{d}x)\nu_t(\mathrm{d}y)\mathrm{d}t \\
&= \iint \langle x - a_t, y - a_t\rangle^3 \nu_t(\mathrm{d}x)\nu_t(\mathrm{d}y)\mathrm{d}t.
\end{aligned}$$

Combine these together, and we complete the proof. □

References

[BL96] D. Bakry and M. Ledoux. Lévy-Gromov's isoperimetric inequality for an infinite dimensional diffusion generator. *Inventiones mathematicae*, 123(1):259–281, December 1996.

[Che21] Yuansi Chen. An Almost Constant Lower Bound of the Isoperimetric Coefficient in the KLS Conjecture. *Geometric and Functional Analysis*, 31(1):34–61, February 2021.

[EAM22] Ahmed El Alaoui and Andrea Montanari. An information-theoretic view of stochastic localization. *IEEE Transactions on Information Theory*, 68(11):7423–7426, November 2022.

[Eld13] Ronen Eldan. Thin Shell Implies Spectral Gap Up to Polylog via a Stochastic Localization Scheme. *Geometric and Functional Analysis*, 23(2):532–569, April 2013.

[Eld23] Ronen Eldan. Analysis of high-dimensional distributions using pathwise methods. pages 4246–4270. December 2023.

[KL22] Bo'az Klartag and Joseph Lehec. Bourgain's slicing problem and KLS isoperimetry up to polylog. *Geometric and Functional Analysis*, 32(5):1134–1159, October 2022.

[KLS95] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13(3-4):541–559, June 1995.

[LV07] László Lovász and Santosh Vempala. The geometry of logconcave functions and sampling algorithms. *Random Structures & Algorithms*, 30(3):307–358, 2007. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/rsa.20135.

[LV18] Yin Tat Lee and Santosh S. Vempala. The Kannan-Lov\'asz-Simonovits Conjecture, July 2018. arXiv:1807.03465 [cs, math].

[LV19] Yin Tat Lee and Santosh S. Vempala. Eldan's Stochastic Localization and the KLS Conjecture: Isoperimetry, Concentration and Mixing, January 2019. arXiv:1612.01507 [cs, math].

[Mil09] Emanuel Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Inventiones mathematicae*, 177(1):1–43, July 2009.

[Mon08] Andrea Montanari. Estimating random variables from random sparse observations. *European Transactions on Telecommunications*, 19(4):385–403, 2008.