WIP Title

Kexing Ying

December 21, 2022

Contents

1 Introduction **3**

1.1 Structure of this essay 3

2 The KLS and Thin-Shell Conjecture **4**

2.1 Concentration 4

2.2 Example: concentration of the Gaussian 5

2.3 The KLS and thin-shell conjecture 5

2.3.1 Equivalent formulation of the KLS conjecture 6

3 The Stochastic Localisation Scheme **8**

3.1 Construction and basic properties 8

3.2 Reduction of KLS to thin-shell 11

3.2.1 Differential of the quadratic variation 11

3.2.2 Analysis of the covariance matrix 12

3.2.3 Stopping the process early 16

4 Almost Constant Bound **18**

1 Introduction

1.1 Structure of this essay

2 The KLS and Thin-Shell Conjecture

2.1 Concentration

Definition 2.1 (Concentration, [Eld18]). Let μ be a measure on $\mathbb{R}^n$, then μ is said to be C-(inversely)-concentrated if for all 1-Lipschitz function $\phi : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{Var}_\mu[\phi] = \mathrm{Var}_{X\sim\mu}[\phi(X)] \le C^2. \tag{1}$$

We denote the least possible such C by C^μ_{con}.

Heuristically, the concentration measures the relation between μ and the Euclidean metric by providing a numerical control for the variance of its norm. This is perhaps best illustrated by the following proposition.

Proposition 2.1. Let X be a $\mathbb{R}^n$-valued random variable. Then for all K-Lipschitz function ϕ : $\mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{Var}[\phi(X)] \le K^2\mathrm{Var}[\|X\|^2].$$

Proof. WLOG. by subtracting its expectation from X, we may assume $\mathbb{E}[X] = 0$. Let X' be a i.i.d. copy of X on the same probability space. Then for all K-Lipschitz function ϕ, we have

$$\begin{aligned}
2\mathrm{Var}[\phi(X)] &= \mathrm{Var}[\phi(X) - \phi(X')] && \text{(i.i.d.)}\\
&= \mathbb{E}[(\phi(X) - \phi(X'))^2] - \mathbb{E}[\phi(X) - \phi(X')]^2\\
&= \mathbb{E}[(\phi(X) - \phi(X'))^2] && \text{(identically distributed)}\\
&\le K^2\mathbb{E}[\|X - X'\|^2] && \text{(as } \phi \text{ is } K\text{-Lipschitz)}\\
&= K^2\mathbb{E}[X^TX + X'^TX' - X^TX' - X'^TX]\\
&= 2K^2\mathrm{Var}[\|X\|^2] - 2K^2\mathrm{Cov}(X,X') = 2K^2\mathrm{Var}[\|X\|^2]. && \text{(independence)}
\end{aligned}$$

implying $\mathrm{Var}[\phi(X)] \le K^2\mathrm{Var}[\|X\|^2]$ as claimed. □

With this proposition in mind, it is clear that for $\mathbb{R}$-valued random variables X, its law μ has concentration $C^\mu_{\text{con}} = \mathrm{Var}[X]$. Furthermore, by considering the projection maps, it follows that the standard Gaussian measure on $\mathbb{R}^n$ is 1-concentrated.

We note that the definition we are presenting here is slightly non-standard. However, utilising a remarkable result due to Milman, we show that this definition is equivalent to the following definitions in a specific sense.

Definition 2.2 (Exponential concentration, [Mil18]). Given a measure μ on $\mathbb{R}^n$, we say μ has exponential concentration if there exists some $c, D > 0$ such that for all 1-Lipschitz function $\phi : \mathbb{R}^n \to \mathbb{R}, t > 0$, we have

$$\mu(|\phi - \mathbb{E}_\mu[\phi]| \ge t) \le ce^{-Dt}. \tag{2}$$

Fixing $c = 1$, we denote the largest possible D as D^μ_{exp}.

Definition 2.3 (First-moment concentration, [Mil18]). Again, given μ a measure on $\mathbb{R}^n$, we say μ has first-moment concentration if there exists some $D > 0$ such that for all 1-Lipschitz function $\phi : \mathbb{R}^n \to \mathbb{R}$, we have

$$\mathbb{E}_\mu[|\phi - \mathbb{E}_\mu[\phi]|] \le \frac{1}{D}. \tag{3}$$

We denote the largest possible D by D^μ_{FM}.

It is clear that exponential concentration implies first-moment concentration. Indeed, if μ has exponential concentration with constant D (taking $c = 1$), then by the tail probability formula,

$$\mathbb{E}_\mu[|\phi - \mathbb{E}_\mu[\phi]|] = \int_0^\infty \mu(|\phi - \mathbb{E}_\mu[\phi]| \geq t)\mathrm{d}t \leq \int_0^\infty e^{-Dt}\mathrm{d}t = \frac{1}{D}.$$

On the other hand, Milman showed that for log-concave measures (namely, measures of the form $\mathrm{d}\mu = \exp(-H)\mathrm{dLeb}^n$ for some convex function $H : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$) on $\mathbb{R}^n$, exponential concentration and first-moment concentration are equivalent in the following sense.

Theorem 1 (Milman, [Mil08])**.** For all log-concave measure μ on $\mathbb{R}^n$, μ has exponential concentration if and only if μ has first-moment concentration. Furthermore, $D^\mu_{\exp} \simeq D^\mu_{\mathrm{FM}}$ where we write $A \simeq B$ if there exists universal constants $C_1, C_2 > 0$ such that $C_1 A \leq B \leq C_2 A$.

With this theorem in mind, we establish the following correspondence.

Proposition 2.2. For all measures μ on $\mathbb{R}^n$, we have

$$\text{Exponentially concentrated} \implies \text{Concentrated} \implies \text{First-moment concentrated}$$

and $D^\mu_{\exp} \leq \sqrt{2}(C^\mu_{\mathrm{con}})^{-1}$ and $(2C^\mu_{\mathrm{con}})^{-1} \leq D^\mu_{\mathrm{FM}}$. Hence, if μ is log-concave, $D^\mu_{\exp} \simeq D^\mu_{\mathrm{FM}} \simeq (C^\mu_{\mathrm{con}})^{-1}$.

Proof. Assume first that μ is C-concentrated. Then by the Chebyshev inequality, we have

$$\mu(|\phi - \mathbb{E}_\mu[\phi]| \geq t) \leq \frac{1}{t^2}\mathrm{Var}_\mu[\phi] \leq \frac{C^2}{t^2},$$

for all 1-Lipschitz ϕ. Thus, by tail probability,

$$\begin{aligned}
\mathbb{E}_\mu[|\phi - \mathbb{E}_\mu[\phi]|] &= \int_0^\infty \mu(|\phi - \mathbb{E}_\mu[\phi]| \geq t)\mathrm{d}t \\
&\leq \inf_{a>0}\left\{\int_0^a \mu(|\phi - \mathbb{E}_\mu[\phi]| \geq t)\mathrm{d}t + C^2\int_a^\infty \frac{1}{t^2}\mathrm{d}t\right\} \\
&\leq \inf_{a>0}\left\{a + \frac{C^2}{a}\right\} = 2C,
\end{aligned}$$

implying μ is first-moment concentrated with respect to the constant $(2C)^{-1}$.

On the other hand, if μ is exponential concentration with some constant D, then again by the tail probability,

$$\mathrm{Var}_\mu[\phi] = \int_0^\infty \mu((\phi - \mathbb{E}_\mu[\phi])^2 \geq t)\mathrm{d}t \leq \int_0^\infty e^{-D\sqrt{t}}\mathrm{d}t = \frac{2}{D^2}$$

implying μ is $\sqrt{2}D^{-1}$-concentrated. $\square$

2.2 Example: concentration of the Gaussian

2.3 The KLS and thin-shell conjecture

Informally, the KLS conjecture suggests that any log-concave measure on $\mathbb{R}^n$ admits the same concentration as that of the Gaussian measure. However, unlike the Gaussian, as the concentration of measures is not invariant under linear functions, it is clear that the KLS conjecture would not hold without a suitable normalization. This leads us to the following formulation of the KLS conjecture.

Conjecture 1 (Kannan-Lovász-Simonovitz, [Eld18])**.** Denoting $\mathscr{M}^n_{\text{con}}$ the set of all log-concave probability measures μ on $\mathbb{R}^n$ satisfying $\text{Var}_\mu[T] \le 1$ for all 1-Lipschitz linear maps $T : \mathbb{R}^n \to \mathbb{R}$, there exists a *universal* constant C such that for all $\mu \in \mathscr{M}^n_{\text{con}}$, μ is C-concentrated.

We remark that C is universal in the sense that it does not depend on any parameter and in particular is independent of the dimension n.

Conjecture 2 (Thin-shell, [Eld13])**.** Taking $\mathscr{M}^n_{\text{con}}$ as above, there exists a universal constant C such that for all $\mu \in \mathscr{M}^n_{\text{con}}$, we have

$$\sqrt{\text{Var}_\mu[\|\cdot\|]} \le C.$$

As the norm function is 1-Lipschitz, it is *a priori* that the thin-shell conjecture is weaker than that of the KLS conjecture. On the other hand, as we shall describe in the next section, as a consequence of the stochastic localisation scheme, Eldan [Eld13] provides a reduction of the KLS conjecture to the thin-shell conjecture up to logarithmic factors.

Theorem 2 (Eldan, [Eld13])**.** Denoting $\mathscr{M}^n_{\text{con}}$ as above, we define

$$C^n_{\text{con}} := \inf\left\{C \mid \forall \mu \in \mathscr{M}^n_{\text{con}},\ \mu \text{ is } C\text{-concentrated}\right\},$$

and

$$C^n_{\text{TS}} := \inf\left\{C \mid \forall \mu \in \mathscr{M}^n_{\text{con}},\ \sqrt{\text{Var}_\mu[\|\cdot\|]} \le C\right\} = \sup_{\mu \in \mathscr{M}^n_{\text{con}}} \sqrt{\text{Var}_\mu[\|\cdot\|]},$$

we have,

$$C^n_{\text{TS}} \le C^n_{\text{con}} \lesssim C^n_{\text{TS}} \log n.$$

We remark that while the constants in theorem 2 depends on the dimension n, the KLS conjecture is reduced to the thin-shell conjecture in the sense that it suffices to show $\sup_n C^n_{\text{TS}} \log n < \infty$. Then, the universal bound for the KLS conjecture is obtained by taking the constant $C = \sup_n C^n_{\text{con}} \le \sup_n C^n_{\text{TS}} \log n < \infty$.

2.3.1 Equivalent formulation of the KLS conjecture

While we have formulated the KLS conjecture using the language of concentration, the conjecture itself is originally formulated as an isoperimetric problem. The isoperimetric problem is the problem in finding the set of unit volume with minimum surface area. In the case of the $\mathbb{R}^n$ equipped with the Lebesgue measure, we have known since the ancient Greeks [Bl05] that the solution is the unit ball. With this in mind, it is natural for us to generalize the problem for arbitrary measures.

Definition 2.4 (Minkowski's boundary measure)**.** Given a measure μ on $\mathbb{R}^n$ and a Borel set $A \subseteq \mathbb{R}^n$, the Minkowski's boundary measure of A,

$$\mu^+(\partial A) := \liminf_{\epsilon \downarrow 0} \frac{\mu(A_\epsilon) - \mu(A)}{\epsilon}.$$

where $A_\epsilon := \{x \in \mathbb{R}^n \mid \text{dist}(x, A) \le \epsilon\}$ is the ϵ-thickening of some Borel set A.

The isoperimetric problem for the measure μ then becomes the problem of finding the set A satisfying $\mu(A) = 1$ with minimum $\mu^+(\partial A)$.

Definition 2.5 (Cheeger's inequality, [Mil08])**.** Given a measure μ on $\mathbb{R}^n$, we say μ satisfy Cheeger's inequality if there exists some D such that for all A,

$$\mu(A) \wedge \mu(A^c) \leq D\mu^+(A).$$

We call the largest such D the inverse Cheeger's constant (or the inverse isoperimetric constant) and denote it by D_{C}^{μ}.

With this, the KLS conjecture can be reformulated as the following.

Conjecture 3 (KLS, [Eld13])**.** Denoting $\mathscr{M}_{\mathrm{iso}}^n$ the set of all log-concave probability measures μ on $\mathbb{R}^n$ satisfying $\mathbb{E}_{X\sim\mu}[X] = 0$ and $\mathrm{Cov}_{X\sim\mu}(X) = \mathrm{id}$ (namely μ is isotropic), there exists a *universal* constant D such that for all $\mu \in \mathscr{M}_{\mathrm{iso}}^n$, μ satisfy the Cheeger's inequality with constant D.

Why are the normalisations equivalent?

The equivalence of the reformulation follows by completing theorem 2.2 with two additional equivalences.

Theorem 3 (Milman, [Mil08])**.** For all log-concave measure μ on $\mathbb{R}^n$, the following are equivalent

- μ has exponential concentration with constant $D_{\exp}^{\mu}$.
- μ has first-moment concentration with constant D_{FM}^{μ}.
- μ satisfy the Cheeger's inequality with constant D_{C}^{μ}.
- μ satisfy the Poincaré inequality: there exists some $D > 0$ such that for all differentiable $\phi : \mathbb{R}^n \to \mathbb{R}$ satisfying $\int \phi \, \mathrm{d}\mu = 0$, we have

$$D \cdot \mathrm{Var}_{\mu}[\phi] \leq \int \|\nabla \phi\|^2 \mathrm{d}\mu.$$

  We denote the largest such D by D_{P}^{μ}.

Furthermore, $D_{\exp}^{\mu} \simeq D_{\mathrm{FM}}^{\mu} \simeq D_{\mathrm{P}}^{\mu} \simeq D_{\mathrm{C}}^{\mu}$.

With this theorem and proposition 2.2 in mind, it is clear that the KLS conjecture can be instead formulated with any of these inequalities instead. Thus, the KLS conjecture can also be phrased using the constant provided by the Poincaré inequality.

Conjecture 4 (KLS, [Eld13])**.** Denoting $\mathscr{M}_{\mathrm{iso}}^n$ as above, there exist a *universal* constant D such that for all $\mu \in \mathscr{M}_{\mathrm{iso}}^n$, μ satisfy the Poincaré inequality with constant D.

We remark that isotropic measures satisfy the normalization condition in 1. Indeed, if $T : \mathbb{R}^n \to \mathbb{R}$ is a 1-Lipschitz linear function, i.e. is of the form $v \mapsto w^T v + d$ for some $w \in S^{n-1}$ and $d \in \mathbb{R}$, then we have

$$\mathrm{Var}_{\mu}[T] = \mathrm{Var}_{X\sim\mu}\left[\sum_{i=1}^{n} w_i X_i + d\right] = \sum_{i,j=1}^{n} w_i w_j \mathrm{Cov}_{X\sim\mu}(X_i, X_j) = \sum_{i=1}^{n} w_i^2 = 1,$$

as $\mathrm{Cov}_{X\sim\mu}(X) = \mathrm{id}$.

3 The Stochastic Localisation Scheme

We will in this section provide a description of the stochastic localisation scheme introduced by Eldan [Eld13] and describe its application in reducing (up to a logarithmic factor) the KLS conjecture into the thin-shell conjecture, i.e. we will describe a proof of theorem 2.

As a high level overview, given a measure, the stochastic localization scheme constructs a measure-valued martingale for which the original measure is recovered in the limit. Then, as the concentration of the measure relates to the covariance of said measure, we will stop the martingale before the covariance grows too large. This allows us to analyze the martingale in a more tractable manner. However, as the the sequence is a martingale, some properties are invariant in time and hence allowing us to conclude that these properties also hold for the original measure.

3.1 Construction and basic properties

Definition 3.1 (Barycenter). Given a (probability) measure μ on $\mathbb{R}^n$, we define its barycenter with respect to the function $F : \mathbb{R}^n \to \mathbb{R}$ to be

$$\bar{\mu}(F) := \int_{\mathbb{R}^n} xF(x)\mu(\mathrm{d}x).$$

In the case that $F = \mathrm{id}$, we simply write $\bar{\mu} = \bar{\mu}(F) = \mathbb{E}_{X\sim\mu}[X]$.

Given the above definition, we now define the following construction central to the stochastic localisation scheme. Let $(W_t)_{t\geq 0}$ be a standard Wiener process in $\mathbb{R}^n$, we define the random functions $(F_t)_{t\geq 0}$ to be the solution of the following infinite system of SDEs:

$$F_0 = 1, \mathrm{d}F_t(x) = \langle x - \bar{\mu}(F_t), \mathrm{d}W_t\rangle F_t(x), \tag{4}$$

for all $x \in \mathbb{R}^n$. We shall from this point forward denote the random variables $a_t := \bar{\mu}(F_t)$. Existence and uniqueness of F.

By applying Itô's formula, we make the following useful observation: for all $x \in \mathbb{R}^n$,

$$\mathrm{d}\log F_t(x) = \frac{\mathrm{d}F_t(x)}{F_t(x)} - \frac{\mathrm{d}[F(x)]_t}{2F_t(x)^2} = \langle x - a_t, \mathrm{d}W_t\rangle - \frac{1}{2}\|x - a_t\|^2\mathrm{d}t \tag{5}$$

where the second equality follows by the construction of F. Hence, as $\log F_0(x) = 0$, we observe

$$\begin{aligned}
\log F_t(x) &= \int_0^t \langle x - a_s, \mathrm{d}W_s\rangle - \frac{1}{2}\int_0^t \|x - a_s\|^2\mathrm{d}s \\
&= \left(\langle x, W_t\rangle - \int_0^t \langle a_s, \mathrm{d}W_s\rangle\right) - \left(\frac{t}{2}\|x\|^2 + \frac{1}{2}\int_0^t \|a_s\|^2\mathrm{d}s - \int_0^t \langle x, a_s\rangle\mathrm{d}s\right) \\
&= -\left(\int_0^t \langle a_s, \mathrm{d}W_s\rangle + \frac{1}{2}\|a_s\|^2\mathrm{d}s\right) + \langle x, a_t + W_t\rangle - \frac{t}{2}\|x\|^2.
\end{aligned}$$

Thus, taking $\mathrm{d}z_t := \langle a_t, \mathrm{d}W_t\rangle + \frac{1}{2}\|a_t\|^2\mathrm{d}t$ and $v_t := a_t + W_t$, we observe $F_t(x)$ is of the form

$$F_t(x) = e^{z_t + \langle x, v_t\rangle - \frac{t}{2}\|x\|^2}, \tag{6}$$

for given Itô processes $(z_t), (v_t)$. With this formulation of $F_t(x)$ in mind, it follows F_t is non-negative, and so, we may define the measure-valued random variable μ_t such that $\mu_t = F_t \mu$, i.e. they have Radon-Nikodym derivative $\mathrm{d}\mu_t / \mathrm{d}\mu = F_t$.

Is $e^{z_t + \langle x, v_t \rangle}$ itself a Itô process?

Proposition 3.1. For all $t \geq 0$, μ_t is a probability measure almost everywhere (a.e.), i.e. $\mathbb{P}(\mu_t(\mathbb{R}^n) = 1) = 1$.

Proof. As μ is a probability measure, it suffices to show $\partial_t \mu_t(\mathbb{R}^n) = 0$. To prove this, we first consider the discrete stochastic integral on the lattice $\Lambda = \mathbb{Z}^n \mathrm{d}t$ for some $\mathrm{d}t > 0$. Then, constructing $\mu_t^{\mathrm{d}t}$ on Λ via the same process as μ_t, for all $t = k\mathrm{d}t \in \Lambda$,

$$\begin{aligned}
\partial_t \mu_t^{\mathrm{d}t}(\mathbb{R}^n) &= \int_{\mathbb{R}^n} \mathrm{d}F_t(x)\mu(\mathrm{d}x) = \int_{\mathbb{R}^n} \langle x - a_t, \mathrm{d}W_t \rangle F_t(x)\mu(\mathrm{d}x) \\
&= \left\langle \int_{\mathbb{R}^n} (x - a_t)F_t(x)\mu(\mathrm{d}x), \mathrm{d}W_t \right\rangle = \langle a_t - a_t\mu_t(\mathbb{R}^n), \mathrm{d}W_t \rangle = 0
\end{aligned}$$

by induction on k. However, by the very construction of the stochastic integral, the densities of $\mu_t^{\mathrm{d}t}$: $F_t^{\mathrm{d}t}$ converges a.e. to F_t as $\mathrm{d}t \to 0$ implying $\mu_t^{\mathrm{d}t} \to \mu_t$ weakly and so, $\mu_t^{\mathrm{d}t}(\mathbb{R}^n) \to \mu_t(\mathbb{R}^n)$ resulting in $\mu_t(\mathbb{R}^n) = 1$ as required. □

We would also like to study the limiting behavior of (μ_t) as $t \to \infty$. To achieve this, we will consider the covariance matrices

$$A_t := \mathsf{Cov}[\mu_t] = \int (x - a_t) \otimes (x - a_t)\mu_t(\mathrm{d}x), \tag{7}$$

where $\otimes$ denotes the Kronecker product. In particular, we will show $(A_t)_{ij} \to 0$ for all $i, j \in \{1, \cdots, n\}$ as $t \to \infty$ allowing us to conclude (μ_t) converges weakly to some Dirac measure. Indeed, this is a direct consequence of the following lemma.

Lemma 3.2 (Brascamp-Lieb inequality, [BL76])**.** Given $V : \mathbb{R}^n \to \mathbb{R}$ convex and $K > 0$, if ν is an isotropic probability measure on $\mathbb{R}^n$ of the form

$$\mathrm{d}\nu = Z e^{-V(x) - \frac{1}{2K}\|x\|^2} \mathrm{dLeb}^n$$

with Z being the normalization constant, then ν satisfy the Poincaré inequality, i.e. for all differentiable ϕ,

$$K \mathrm{Var}_\nu[\phi] \leq \int \|\nabla \phi\|^2 \mathrm{d}\nu.$$

Proof of lemma 3.2. TODO. Maybe go in appendix? □

With this lemma in mind, by taking $\nu = \mu_t$ using equation (6) and defining $\pi_i(x) := x_i$, we have by the Cauchy-Schwarz inequality

$$(A_t)_{ij} \leq \sqrt{\mathrm{Var}_{\mu_t}[\pi_i]}\sqrt{\mathrm{Var}_{\mu_t}[\pi_j]} \leq \max_{k=1,\cdots,n} \frac{1}{t} \int \|\nabla \pi_k\|^2 \mathrm{d}\mu_t$$

Again, using equation (6), we note that any realizations of $(F_t(x))$ is eventually decreasing in t for all $x \neq 0$, implying

$$\sup_{t>0} \max_{k=1,\cdots,n} \int \|\nabla \pi_k\|^2 \mathrm{d}\mu_t = \sup_{t>0} \max_{k=1,\cdots,n} \int x_k^2 \mathrm{d}\mu_t < \infty.$$

Thus, by taking $t \to \infty$ we have $(A_t)_{ij} \to 0$ for all $i, j \in \{1, \cdots, n\}$ as claimed and we have the following corollary.

Corollary 3.3. (μ_t) converges weakly to some Dirac measure almost everywhere. We denote this limiting (random) Dirac measure by δ_{a_∞} where a_∞ is some $\mathbb{R}^n$-valued random variable.

Since convergence implies relatively compact, applying the Dunford-Pettis theorem it follows that any realizations of (F_t) is uniformly integrable. Thus, we can make the following deductions about a_∞.

Corollary 3.4. The massive point a_∞ of the limiting Dirac measure is the limit of a_t as $t \to \infty$ and has law μ.

Proof. Since (F_t) is uniformly integrable we have convergence of means almost everywhere, namely

$$a_t = \int x \mu_t(\mathrm{d}x) \xrightarrow{\text{a.e.}} \int x \delta_{a_\infty}(\mathrm{d}x) =: a_\infty \text{ as } t \to \infty$$

implying that a_t converges a.e. to a_∞ as $t \to \infty$ as required.

Furthermore, taking $\phi : \mathbb{R}^n \to \mathbb{R}$ to be any bounded continuous function, we have

$$\int \phi(x)\mu(\mathrm{d}x) = \lim_{t\to\infty} \int \phi(x)\mu_t(\mathrm{d}x).$$

Then, taking expectation on both sides, we obtain

$$\begin{aligned}
\int \phi(x)\mu(\mathrm{d}x) &= \mathbb{E}\left[\lim_{t\to\infty} \int \phi(x)\mu_t(\mathrm{d}x)\right] \\
&= \lim_{t\to\infty} \mathbb{E}\left[\int \phi(x)\mu_t(\mathrm{d}x)\right] && \text{(Dominated convergence)} \\
&= \lim_{t\to\infty} \mathbb{E}[\phi(a_t)] && \text{(LOTUS. theorem)} \\
&= \mathbb{E}[\phi(a_\infty)]. && \text{(Dominated convergence \& continuity of } \phi)
\end{aligned}$$

Thus, $\mathbb{E}_\mu[\phi] = \mathbb{E}[\phi(a_\infty)]$ for all bounded continuous ϕ implying $a_\infty \sim \mu$. □

Proposition 3.5. For all $x \in \mathbb{R}^n$, $(F_t(x))_{t\geq 0}$ is a martingale. Furthermore, for any continuous $\phi : \mathbb{R}^n \to \mathbb{R}$, defining the process $M_t := \int \phi \mathrm{d}\mu_t$, $(M_t)_{t\geq 0}$ is also a martingale.

Proof. By its very construction, $(F_t(x))$ is a martingale by observing equation 4 has no drift term. Now, for all $s \leq t$ we have by the conditional Fubini's theorem,

$$\begin{aligned}
\mathbb{E}[M_t \mid \mathscr{F}_s] &= \mathbb{E}\left[\int \phi(x)F_t(x)\mu(\mathrm{d}x) \middle| \mathscr{F}_s\right] \\
&= \int \phi(x)\mathbb{E}[F_t(x) \mid \mathscr{F}_s]\mu(\mathrm{d}x) = \int \phi(x)F_s(x)\mu(\mathrm{d}x) = M_s
\end{aligned}$$

implying (M_t) is also a martingale. □

Using the same proof as corollary 3.4, we observe

$$M_t \xrightarrow{\text{a.e.}} M_\infty \sim \phi_*\mu \tag{8}$$

where $\phi_*\mu$ denotes the push-forward measure of μ along ϕ.

3.2 Reduction of KLS to thin-shell

In this section we will present a proof of theorem 2 as presented in [Eld18] based on the method as described by [LV16]. Furthermore, by making small modifications to this method, we will also present the main result of [LV16].

We recall the goal of theorem 2 is to control $\mathrm{Var}_\mu[\phi]$ by a logarithmic factor of $\mathrm{Var}_\mu[\|\cdot\|]$. As translating the barycenter of μ does not affect its variance, we may assume μ has its barycenter $\overline{\mu}$ at the origin. Furthermore, we may assume μ is supported on $B_n(0) \subseteq \mathbb{R}^n$ with $B_n(0)$ the ball at the origin of radius n. Thus, we also have explain

$$\operatorname{supp} \mu_t = \operatorname{supp} F_t\mu \subseteq \operatorname{supp} \mu \subseteq B_n(0)$$

for all $t > 0$.

Fix $\phi : \mathbb{R}^n \to \mathbb{R}$ some 1-Lipschitz function and let (M_t) be the martingale as described above and in particular we recall equation (3.4) and so, $\mathrm{Var}_\mu[\phi] = \mathrm{Var}[M_\infty]$ where $M_t \xrightarrow{\text{a.e.}} M_\infty$. Hence, for all $t > 0$, by the martingale property we have

$$\begin{aligned}\mathrm{Var}[M_t] + \mathbb{E}[\mathrm{Var}[M_\infty \mid \mathscr{F}_t]] &= (\mathbb{E}[M_t^2] - \mathbb{E}[M_t]^2) + \mathbb{E}\left[\mathbb{E}[M_\infty^2 \mid \mathscr{F}_t] - \mathbb{E}[M_\infty \mid \mathscr{F}_t]^2\right]\\ &= \mathbb{E}[M_t^2] + (\mathbb{E}[\mathbb{E}[M_\infty^2 \mid \mathscr{F}_t]] - \mathbb{E}[M_t^2])\\ &= \mathbb{E}[M_\infty^2] = \mathrm{Var}[M_\infty],\end{aligned}$$

where the second equality follows as $\mathbb{E}[M_t] = \mathbb{E}[M_\infty] = \mathbb{E}_\mu[\phi] = 0$ as a linear map

On the other hand, as (M_t) is a martingale, $M_t^2 - [M]_t$ is also a martingale implying $\mathbb{E}[M_t^2] = \mathbb{E}[M]_t$ and so $\mathrm{Var}[M_t] = \mathbb{E}[M_t^2] - \mathbb{E}[M_t]^2 = \mathbb{E}[M]_t - \overline{\mu}^2 = \mathbb{E}[M]_t$. Hence, combining this with the above, we obtain the bound

$$\mathrm{Var}_\mu[\phi] = \mathrm{Var}[M_\infty] = \mathrm{Var}[M]_t + \mathbb{E}[\mathrm{Var}[M_\infty \mid \mathscr{F}_t]] = \mathbb{E}[M]_t + \mathbb{E}[\mathrm{Var}_{\mu_t}[\phi]]. \tag{9}$$

Now, observing that ϕ is 1-Lipschitz implies $\|\nabla\phi\|^2 \leq 1$, we have by lemma 3.2 the bound $\mathrm{Var}_{\mu_t}[\phi] \leq t^{-1}$ (in fact $\mathrm{Var}_{\mu_t}[\phi] \leq t^{-1} \wedge n^2$ as we have assumed $\operatorname{supp} \mu_t \subseteq B_n(0)$). Thus, the second term $\mathbb{E}[\mathrm{Var}_{\mu_t}[\phi]]$ is bounded by t^{-1}. With this in mind, by choosing an appropriate random time τ to stop the process such that $\mathbb{E}[M]_\tau$ is nicely bounded, the result follow by bounding $\mathbb{E}[\tau^{-1}]$. We dedicate the remainder of this section to describe said procedure in detail.

3.2.1 Differential of the quadratic variation

To bound the term $\mathbb{E}[M]_\tau$ we will compute its differential and bound it sufficiently such that we reobtain a bound for $[M]_\tau$ after integration. We will show $\mathrm{d}[M]_t$ is bounded by a quantity concerning A_t. This should not be at all surprising as both $\mathrm{d}[M]_t$ and A_t describes the variation of M_t in a infinitesimal time neighborhood of t.

We compute

$$\begin{aligned}
\mathrm{d}M_t &= d\int \phi(x)F_t(x)\mu(\mathrm{d}x) = \int \phi(x)\langle x-a_t, \mathrm{d}W_t\rangle \mu_t(\mathrm{d}x) \\
&= \left\langle \int \phi(x)(x-a_t)\mu_t(\mathrm{d}x), \mathrm{d}W_t \right\rangle
\end{aligned}$$

and so, by considering the component-wise quadratic variation, we have

$$\mathrm{d}[M]_t = \left\| \int \phi(x)(x-a_t)\mu_t(\mathrm{d}x) \right\|^2 \mathrm{d}t. \tag{10}$$

Then, denoting θ the vector $\int \phi(x)(x-a_t)\mu_t(\mathrm{d}x)$ normalized to have norm 1, so

$$\left\langle \theta, \int \phi(x)(x-a_t)\mu_t(\mathrm{d}x) \right\rangle = \left\| \int \phi(x)(x-a_t)\mu_t(\mathrm{d}x) \right\|$$

we observe,

$$\begin{aligned}
\mathrm{d}[M]_t &= \left\langle \theta, \int \phi(x)(x-a_t)\mu_t(\mathrm{d}x) \right\rangle^2 \mathrm{d}t = \left\langle \theta, \int (\phi(x)-a_t)(x-a_t)\mu_t(\mathrm{d}x) \right\rangle^2 \mathrm{d}t \\
&= \left(\int (\phi(x)-a_t)\langle \theta, x-a_t\rangle \mu_t(\mathrm{d}x) \right)^2 \mathrm{d}t \\
&\leq \left(\int (\phi(x)-a_t)^2 \mu_t(\mathrm{d}x) \right)\left(\int \langle \theta, x-a_t\rangle^2 \mu_t(\mathrm{d}x) \right) \mathrm{d}t \\
&= \mathrm{Var}_{\mu_t}[\phi]\left(\int \theta^T (x-a_t)^{\otimes 2}\theta \mu_t(\mathrm{d}x) \right)\mathrm{d}t = \mathrm{Var}_{\mu_t}[\phi](\theta^T A_t \theta)\mathrm{d}t \\
&\leq \mathrm{Var}_{\mu_t}[\phi]\|A_t\|_{\mathrm{op}}\mathrm{d}t.
\end{aligned} \tag{11}$$

where the inequality follows by the Cauchy-Schwarz inequality and $\|\cdot\|_{\mathrm{op}}$ denotes the operator norm. Thus, as we know $\mathrm{Var}_{\mu_t}[\phi] \leq t^{-1}$, the problem is now reduced to that of bounding $\|A_t\|_{\mathrm{op}}$.

3.2.2 Analysis of the covariance matrix

As demonstrated in section 3.1, we know the limiting behavior of the covariance matrices, namely $A_t \to 0$ point-wise as $t \to \infty$. This was important for us to establish the existence of the limit of (a_t) and (M_t). However, as shown above, we now require some quantitative bounds for the operator norm of A_t. For this purpose, we first compute some useful properties of A_t.

Observing

$$\int \mathrm{d}F_t(x)\mu(\mathrm{d}x) = \int \langle x-a_t, \mathrm{d}W_t\rangle \mu_t(\mathrm{d}x) = \left\langle \int x\mu_t(\mathrm{d}x) - a_t, \mathrm{d}W_t \right\rangle = 0,$$

we have

$$\begin{aligned}
\mathrm{d}a_t &= \mathrm{d}\int xF_t(x)\mu(\mathrm{d}x) = \int x\mathrm{d}F_t(x)\mu(\mathrm{d}x) = \int (x-a_t)\mathrm{d}F_t(x)\mu(\mathrm{d}x) \\
&= \int (x-a_t)\langle x-a_t, \mathrm{d}W_t\rangle F_t(x)\mu(\mathrm{d}x) = \int (x-a_t)^{\otimes 2}\mathrm{d}W_t \mu_t(\mathrm{d}x) = A_t \mathrm{d}W_t
\end{aligned} \tag{12}$$

where the second to last equality used the fact that $v\langle v,w\rangle = v^{\otimes 2}w$ for any appropriate v,w. Similarly, computing using Itô's formula, we have

$$\begin{aligned} \mathrm{d}A_t &= \mathrm{d}\int (x-a_t)^{\otimes 2}F_t(x)\mu(\mathrm{d}x) \\ &= \int (x-a_t)^{\otimes 2}\mathrm{d}F_t(x) + F_t(x)\,\mathrm{d}(x-a_t)^{\otimes 2} \\ &\qquad - 2(x-a_t)\otimes \mathrm{d}[a_t,F_t(x)]_t + F_t(x)\mathrm{d}[a_t]_t\mu(\mathrm{d}x). \end{aligned} \tag{13}$$

The second term vanishes as

$$\int F_t(x)\,\mathrm{d}(x-a_t)^{\otimes 2}\,\mu(\mathrm{d}x) = -2\mathrm{d}a_t \otimes \overbrace{\int (x-a_t)\mu_t(\mathrm{d}x)}^{=0} = 0.$$

Also, by equation (12), $\mathrm{d}a_t = A_t\mathrm{d}W_t$ implying $\mathrm{d}[a_t]_t = A_t^2\mathrm{d}t$. Finally, as both (a_t) and $(F_t(x))$ are martingales, $\mathrm{d}[a_t,F_t(x)]_t = F_t(x)A_t x\mathrm{d}t$ and the third term becomes

$$\begin{aligned} -2\int (x-a_t)\otimes \mathrm{d}[a_t,F_t(x)]\mu_t(\mathrm{d}x) &= -2A_t\left(\int (x-a_t)\otimes x\mu_t(\mathrm{d}x)\right)\mathrm{d}t \\ &= -2A_t\left(\overbrace{\int (x-a_t)^{\otimes 2}\mu_t(\mathrm{d}x)}^{A_t} + \overbrace{\int (x-a_t)\mu_t(\mathrm{d}x)}^{=0}\otimes a_t\right)\mathrm{d}t \\ &= -2A_t^2\mathrm{d}t. \end{aligned}$$

Hence, combining these and equation (4) together in (13), we have

$$\mathrm{d}A_t = \int (x-a_t)^{\otimes 2}\langle x-a_t,\mathrm{d}W_t\rangle\mu_t(\mathrm{d}x) - A_t^2\mathrm{d}t$$

However, since we wish to bound A_t from above, as the drift term $-A_t^2\mathrm{d}t$ only contributes negatively, an upper bound for the process of the form $\int (x-a_t)^{\otimes 2}\langle x-a_t,\mathrm{d}W_t\rangle\mu_t(\mathrm{d}x)$ is also sufficient for A_t. Hence, we proceed by ignoring the drift term and redefine the process A_t such that

$$\mathrm{d}A_t = \int (x-a_t)^{\otimes 2}\langle x-a_t,\mathrm{d}W_t\rangle\mu_t(\mathrm{d}x). \tag{14}$$

With this justification, we now proceed to bound the operator norm of this new A_t. In particular, as A_t is symmetric, we recall that $\|A_t\|_{\mathrm{op}} = \max_{i=1,\cdots,n}\lambda_i(t) = \|(\lambda_i(t))_{i=1}^n\|_\infty$ where $\lambda_i(t)$ denotes the distinct eigenvalues of A_t. Hence, it suffices to find a bound for the potential

$$\Phi^\alpha(t) = \sum_{i=1}^n |\lambda_i(t)|^\alpha = \|(\lambda_i(t))_{i=1}^n\|_\alpha^\alpha \tag{15}$$

for some $\alpha > 0$. Furthermore, as A_t is positive semi-definite, $\lambda_i(t) \geq 0$ for all $i = 1,\cdots,n$ and thus we have $\Phi^\alpha(t) = \sum_{i=1}^n \lambda_i(t)^\alpha$. Again, to proceed, we will attempt to compute $\mathrm{d}\Phi^\alpha(t)$ at some $t = t_0 > 0$ utilizing the following lemma.

Lemma 3.6. If $A=[a_{ij}]$ is a diagonal matrix with distinct eigenvalues $\lambda_i,\cdots,\lambda_n$, then for all $i,j,k,l,m\in 1,\cdots n$, we have

- $\frac{\partial\lambda_i}{\partial a_{jk}}=\delta_{ij}\delta_{ik}$;
- whenever $i\neq j$, $\frac{\partial^2\lambda_i}{\partial a_{ij}^2}=2(\lambda_i-\lambda_j)^{-1}$;
- and for $j\neq l,k\neq m$ or $i\neq j$ and $i\neq k$, $\frac{\partial^2\lambda_i}{\partial a_{jk}\partial a_{lm}}=0$,

where δ_{ij} denotes the Kronecker delta function.

As this lemma requires the matrix to be diagonal, denoting $e_1,\cdots,e_n$ as the normalized eigenbasis of A_{t_0} (they are in fact orthonormal as A_{t_0} is positive semi-definite), we will consider A_t with respect to this basis by considering the entries

$$a_{ij}(t):=\langle e_i,A_te_j\rangle.$$

Using equation (14), we compute

$$\begin{aligned}\mathrm{d}a_{ij}(t)&=\left\langle e_i,\left(\int(x-a_t)^{\otimes 2}\langle x-a_t,\mathrm{d}W_t\rangle\mu_t(\mathrm{d}x)\right)e_j\right\rangle\\&=\left\langle\int\langle e_i,(x-a_t)^{\otimes 2}e_j\rangle(x-a_t)\mu_t(\mathrm{d}x),\mathrm{d}W_t\right\rangle=\langle\xi_{ij},\mathrm{d}W_t\rangle\end{aligned}$$

where we introduce the notation $\xi_{ij}=\int\langle e_i,(x-a_t)^{\otimes 2}e_j\rangle(x-a_t)\mu_t(\mathrm{d}x)$. Thus, combining this with lemma 3.6, denoting $\lambda_i=\lambda_i(t_0)$, we have by Itô's formula

$$\begin{aligned}\mathrm{d}\lambda_i(t)&=\sum_{j,k=1}^n\frac{\partial\lambda_i}{\partial a_{jk}}\mathrm{d}a_{jk}(t)+\frac{1}{2}\sum_{j,k=1}^n\sum_{l,m=1}^n\frac{\partial^2\lambda_i}{\partial a_{jk}\partial a_{lm}}\mathrm{d}[a_{jk},a_{lm}]_t\\&=\langle\xi_{ii},\mathrm{d}W_t\rangle+\sum_{j\neq i}\frac{\mathrm{d}[a_{ij}]_t}{\lambda_i-\lambda_j}=\langle\xi_{ii},\mathrm{d}W_t\rangle+\sum_{j\neq i}\frac{\|\xi_{ij}\|^2}{\lambda_i-\lambda_j}\mathrm{d}t.\end{aligned}\tag{16}$$

at $t=t_0$. As a result, it is also clear that $\mathrm{d}[\lambda_i(t)]_{t_0}=\|\xi_{ii}\|^2\mathrm{d}t$.
Again applying Itô's formula, we may finally compute

Move computation to appendix

$$
\begin{aligned}
\mathrm{d}\Phi^\alpha(t) &= \sum_{i=1}^n \frac{\partial \Phi^\alpha}{\partial \lambda_i}\Big|_{t=t_0} \mathrm{d}\lambda_i(t) + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2\Phi^\alpha}{\partial\lambda_i\partial\lambda_j}\Big|_{t=t_0}\mathrm{d}[\lambda_i,\lambda_j]_t \\
&= \alpha\sum_{i=1}^n \lambda_i^{\alpha-1}\mathrm{d}\lambda_i(t) + \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^n \lambda_i^{\alpha-2}\mathrm{d}[\lambda_i(t)]_t \\
&= \alpha\sum_{i=1}^n \lambda_i^{\alpha-1}\left(\langle \xi_{ii}, \mathrm{d}W_t\rangle + \sum_{j\neq i}\frac{\|\xi_{ij}\|^2}{\lambda_i-\lambda_j}\mathrm{d}t\right) + \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^n \lambda_i^{\alpha-2}\mathrm{d}[\lambda_i(t)]_t \\
&= \alpha\sum_{i\neq j}\lambda_i^{\alpha-1}\frac{\|\xi_{ij}\|^2}{\lambda_i-\lambda_j}\mathrm{d}t + \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^n\lambda_i^{\alpha-2}\|\xi_{ii}\|^2\mathrm{d}t + \Big\langle \underbrace{\alpha\sum_{i=1}^n \lambda_i^{\alpha-1}\xi_{ii}}_{=:v_t}, \mathrm{d}W_t\Big\rangle \\
&= \frac{1}{2}\alpha\sum_{i\neq j}\|\xi_{ij}\|^2\frac{\lambda_i^{\alpha-1}-\lambda_j^{\alpha-1}}{\lambda_i-\lambda_j}\mathrm{d}t + \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^n\lambda_i(t)^{\alpha-2}\|\xi_{ii}\|^2\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle \\
&\leq \frac{1}{2}\alpha(\alpha-1)\sum_{i\neq j}\|\xi_{ij}\|^2(\lambda_i\vee\lambda_j)^{\alpha-2}\mathrm{d}t + \frac{1}{2}\alpha(\alpha-1)\sum_{i=1}^n\lambda_i(t)^{\alpha-2}\|\xi_{ii}\|^2\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle \\
&= \frac{1}{2}\alpha(\alpha-1)\sum_{i,j=1}^n\|\xi_{ij}\|^2(\lambda_i\vee\lambda_j)^{\alpha-2}\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle \leq \alpha^2\sum_{i,j=1}^n\|\xi_{ij}\|^2\lambda_i^{\alpha-2}\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle,
\end{aligned}
$$

where the first inequality holds as

$$
\frac{\lambda_i^{\alpha-1}-\lambda_j^{\alpha-1}}{\lambda_i-\lambda_j} = \lambda_i^{\alpha-2} + \lambda_i^{\alpha-3}\lambda_j + \cdots + \lambda_i^{\alpha-2} \leq (\alpha-1)(\lambda_i\vee\lambda_j)^{\alpha-2}.
$$

Thus, we have shown

$$
\mathrm{d}\Phi^\alpha(t) \leq \alpha^2\sum_{i=1}^n\lambda_i(t)^{\alpha-2}\sum_{j=1}^n\|\xi_{ij}\|^2\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle \tag{17}
$$

where $v_t := \alpha\sum_{i=1}^n \lambda_i^{\alpha-1}\xi_{ii}$.

By recalling that our goal is to bound $\|A_t\|_{\mathrm{op}}$ from above (c.f. equation (9) and (11)), we may assume without loss of generality that $\|A_t\|_{\mathrm{op}} \geq 1$. Thus, applying the reverse Cauchy-Schwarz inequality to equation (17), we have

$$
\begin{aligned}
\mathrm{d}\Phi^\alpha(t) &\leq 2\alpha^2\sum_{i=1}^n\lambda_i(t)^{\alpha-2}\sum_{j=1}^n\|\xi_{ij}\|^2\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle \\
&\leq 2\alpha^2\|A_t\|_{\mathrm{op}}^2\sum_{i=1}^n\lambda_i(t)^{\alpha-2}\sum_{j=1}^n\|\xi_{ij}\|^2\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle \\
&\lesssim 2\alpha^2\sum_{i=1}^n\lambda_i(t)^{\alpha}\sum_{j=1}^n\|\xi_{ij}\|^2\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle.
\end{aligned}
$$

Thus, defining $K_t := \sup_i \sum_{j=1}^n \|\xi_{ij}\|^2$, we have the bound

$$
\mathrm{d}\Phi^\alpha(t) \lesssim 2\alpha^2 K_t\Phi^\alpha(t)\mathrm{d}t + \langle v_t, \mathrm{d}W_t\rangle. \tag{18}
$$

3.2.3 Stopping the process early

As outlined in the beginning of this section, we will stop the process early in order to provide a bound for the right hand side of equation (9). By observing equation (11), we hypothesize that we should stop the process once $\|A_t\|_{\mathrm{op}}$ grows too large. As a result we define the stopping time

$$\tau := \inf\{t > 0 \mid \|A_t\|_{\mathrm{op}} > 2\} \wedge 1.$$

By the optional stopping theorem we have

$$\begin{aligned}[M]_\tau = \int_0^\tau \mathrm{d}[M]_t &\leq \int_0^\tau \overbrace{\mathrm{Var}_\mu[\phi]}^{\leq t^{-1}\wedge n^2} \overbrace{\|A_t\|_{\mathrm{op}}}^{\leq 2} \mathrm{d}t \\ &\leq 2\int_0^\tau t^{-1} \wedge n^2 \mathrm{d}t \leq 2\int_0^1 t^{-1} \wedge n^2 \mathrm{d}t = 2 + 4\log n.\end{aligned}$$

Combining this with equation (9), we obtain

$$\mathrm{Var}_\mu[\phi] \leq 2 + 4\log n + \mathbb{E}[\tau^{-1}], \tag{19}$$

and it remains to find an upper bound for $\mathbb{E}[\tau^{-1}]$. Observing that $t < \tau$ whenever $\Phi^\alpha(t) < 2^\alpha$, we define the σ the first time for which the potential $\Phi^\alpha(t)$ reaches 2^α, namely

$$\sigma := \inf\{t > 0 \mid \Phi^\alpha(t) = 2^\alpha\},$$

we have $\sigma^{-1} \geq \tau^{-1}$ and so it suffices to bound σ from below.

For simplicity, let us ignore the stochastic term in equation (18) and regard it as an ODE. Then, by Gronwall's inequality, if we can find some constant K such that $K_t \leq K$ for all $t \leq \tau$, we have the bound

$$S_t \leq ne^{2\alpha^2 K t}.$$

try to not ignore the martingale term

Thus, substituting σ into the above, we have

$$2^\alpha = S_\sigma \leq ne^{2\alpha^2 K\sigma}$$

implying

$$\frac{\alpha\log 2 - \log n}{2\alpha^2 K} \leq \sigma \leq \tau.$$

Then, taking $\alpha = 10K\log n$, it is easy to check that

$$\frac{1}{10K\log n} \leq \frac{\alpha\log 2 - \log n}{2\alpha^2 K}$$

implying $\mathbb{E}[\tau^{-1}] \leq 10K\log n$. Of course, this deduction only holds while ignoring the stochastic term $\langle v_t, \mathrm{d}W_t\rangle$. Nonetheless, this is justified as one can show that $\|v_t\|_2$ is bounded $\alpha\Phi^\alpha(t)$ and so the same analysis holds by applying the stochastic Gronwall's inequality (c.f. second part of lemma 34 in [LV18]).

Finally, to find a bound for (K_t), we employ the following lemma.

Lemma 3.7 (Lemma 1.6 in [Eld13])**.** Denoting C^n_{TS} as in theorem 2, there exists a constant C such that for any log-concave, isotropic probability measure μ, we have

$$\sup_{\theta\in S^{n-1}}\sum_{i,j=1}^{n}\mathbb{E}_{X\sim\mu}[\langle X,e_i\rangle\langle X,e_j\rangle\langle X,\theta\rangle]^2\le C\sum_{k=1}^{n}\frac{(C^n_{\mathrm{TS}})^2}{k},$$

where $\{e_1,\cdots,e_n\}$ is any orthonormal basis on $\mathbb{R}^n$.

Recalling that

$$\xi_{ij}=\mathbb{E}_{X+a_t\sim\mu_t}[\langle e_i,X^{\otimes 2}e_j\rangle X]=\mathbb{E}_{X+a_t\sim\mu_t}[\langle X,e_i\rangle\langle X,e_j\rangle X],$$

we have by Parseval's identity

$$\begin{aligned}
K_t&=\sup_i\sum_{j=1}^{n}\|\xi_{ij}\|^2=\sup_i\sum_{j=1}^{n}\|\mathbb{E}_{X+a_t\sim\mu_t}[\langle X,e_i\rangle\langle X,e_j\rangle X]\|^2\\
&=\sup_i\sum_{j=1}^{n}\sum_{k=1}^{n}\left\langle\mathbb{E}_{X+a_t\sim\mu_t}[\langle X,e_i\rangle\langle X,e_j\rangle X],e_k\right\rangle^2\\
&=\sup_i\sum_{j,k=1}^{n}\mathbb{E}_{X+a_t\sim\mu_t}[\langle X,e_i\rangle\langle X,e_j\rangle\langle X,e_k\rangle]^2\\
&\le\sup_{\theta\in S^{n-1}}\sum_{i,j=1}^{n}\mathbb{E}_{X+a_t\sim\mu}[\langle X,e_i\rangle\langle X,e_j\rangle\langle X,\theta\rangle]^2.
\end{aligned}$$

We note that we cannot direct apply lemma 3.7 at this point since the measure μ_t might not be isotropic. Hence, to be able to use the lemma, we need to normalize the covariance of μ_t. Namely, taking $X+a_t\sim\mu_t$, we define $Y=A^{-1/2}X$ which by construction is isotropic. Thus, by observing that

$$\mathbb{E}_{X+a_t\sim\mu}[\langle X,e_i\rangle\langle X,e_j\rangle\langle X,\theta\rangle]^2\le\|A_t\|^3_{\mathrm{op}}\mathbb{E}_{X+a_t\sim\mu}[\langle Y,e_i\rangle\langle Y,e_j\rangle\langle Y,\theta\rangle]^2,$$

we have

$$\begin{aligned}
K_t&\le\sup_{\theta\in S^{n-1}}\sum_{i,j=1}^{n}\mathbb{E}_{X+a_t\sim\mu}[\langle X,e_i\rangle\langle X,e_j\rangle\langle X,\theta\rangle]^2\\
&\le\|A_t\|^3_{\mathrm{op}}\sup_{\theta\in S^{n-1}}\sum_{i,j=1}^{n}\mathbb{E}_{X+a_t\sim\mu}[\langle Y,e_i\rangle\langle Y,e_j\rangle\langle Y,\theta\rangle]^2\le 8C\sum_{k=1}^{n}\frac{(C^n_{\mathrm{TS}})^2}{k}
\end{aligned}\tag{20}$$

where the last inequality follows as $\|A_t\|_{\mathrm{op}}\le 2$ for all $t<\tau$.

At last, combining equation (20) and (19), we have

$$\mathrm{Var}_\mu[\phi]\le 2+\log n\left(4+\overbrace{80C\sum_{k=1}^{n}\frac{1}{k}(C^n_{\mathrm{TS}})^2}^{\Theta(\log n)}\right)=\Theta_n((C^n_{\mathrm{TS}}\log n)^2)$$

implying there exists a constant $R>0$ such that for all 1-Lipschitz ϕ, $\sqrt{\mathrm{Var}_\mu[\phi]}\le RC^n_{\mathrm{TS}}\log n$, i.e. μ is $RC^n_{\mathrm{TS}}\log n$-concentrated and so, $C^n_{\mathrm{con}}\le RC^n_{\mathrm{TS}}\log n$ as required.

4 Almost Constant Bound

References

[BL76] Herm Jan Brascamp and Elliott H Lieb. On extensions of the brunn-minkowski and prékopa-leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4):366–389, 1976.

[Bl05] Viktor Blåsjö. The isoperimetric problem. *The American Mathematical Monthly*, 112(6):526–566, 2005.

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

[Eld18] Ronen Eldan. Lecture notes - from stochastic calculus to geometric inequalities. 2018.

[LV16] Yin Tat Lee and Santosh S. Vempala. Eldan's stochastic localization and the kls conjecture: Isoperimetry, concentration and mixing, 2016.

[LV18] Yin Tat Lee and Santosh S. Vempala. The kannan-lovász-simonovits conjecture, 2018.

[Mil08] Emanuel Milman. Uniform tail-decay of lipschitz functions implies cheeger's isoperimetric inequality under convexity assumptions. *Comptes Rendus Mathematique*, 346(17):989–994, 2008.

[Mil18] Emanuel Milman. On the role of convexity in isoperimetry, spectral-gap and concentration. 2018.