
THE KLS CONJECTURE

Amit Rajaraman

Last updated June 12, 2021

§0. Notation

- We refer to measures by greek symbols such as μ and ν and their densities by lowercase alphabets beginning from p .
- B refers to the Euclidean ball of radius 1 in \mathbb{R}^n (the value of n is usually understandable from context).
- Given a measure μ on \mathbb{R}^n and an $(n-1)$ -dimensional surface(?) S in \mathbb{R}^n , $\mu^+(S)$ refers to the “surface area” of the set S , that is,

$$\mu^+(X) = \lim_{\varepsilon \rightarrow 0} \frac{\mu(X + \varepsilon B)}{2\varepsilon}.$$

Alternatively, if $X \subseteq \mathbb{R}^n$ is compact, then

$$\mu^+(\partial X) = \lim_{\varepsilon \rightarrow 0} \frac{\mu(X + \varepsilon B) - \mu(X)}{\varepsilon}.$$

- While needles in [KLS95] refer to one-dimensional segments with a polynomial factor (ℓ^{n-1} where ℓ is linear) in particular, we use them more generally to refer to log-concave measures with a one-dimensional support.

§1. Measure Disintegration

1.1. Introduction

To begin with, let us describe a needle decomposition procedure given in [LV18] to prove the KLS Conjecture. Suppose that we are given a log-concave measure μ with density p with compact convex support K . Let us also fix a subset $E \subseteq K$ of measure $1/2$. We would like to bound $\mu^+(\partial E)$ below (over all such E).

Now, suppose that we have some hyperplane H that divides space into two half-spaces H_1 and H_2 . Let $K_i = K \cap H_i$ and further assume that $\mu(E \cap H_i) = \frac{1}{2}\mu(K_i)$ for each i . Consider the measures μ_1 and μ_2 with densities given by

$$p_i(x) = \begin{cases} p(x) \frac{\mu(K)}{\mu(K_i)}, & x \in K_i, \\ 0, & \text{otherwise.} \end{cases}$$

Observe that

$$\begin{aligned} p &= p_1 \frac{\mu(K_1)}{\mu(K)} + p_2 \frac{\mu(K_2)}{\mu(K)} \\ \mu &= \mu_1 \frac{\mu(K_1)}{\mu(K)} + \mu_2 \frac{\mu(K_2)}{\mu(K)} \end{aligned} \tag{1.1}$$

More generally, suppose we have some space Ω with a probability measure ν on it such that

$$\mu = \int_{\Omega} \mu_{\omega} d\nu(\omega), \tag{1.2}$$

where the (μ_ω) are log-concave measures on \mathbb{R}^n . In the above example, we can take $\Omega = \{1, 2\}$ and $\nu(\{i\}) = \mu(K_i)/\mu(K)$ for $i \in \Omega$.

Then, given any set E of measure $1/2$, we have

$$\begin{aligned} \mu^+(\partial E) &= \int_{\Omega} \mu_\omega^+(\partial E) d\nu(\omega) \\ &\geq \int_{\Omega} \psi_\omega \mu_\omega(E) (1 - \mu_\omega(E)) d\nu(\omega), \end{aligned} \quad (1.3)$$

where ψ_ω is the isoperimetric constant of μ_ω . If we manage to bound the expression on the right below by some constant independent of E , then the KLS conjecture follows. It is also worth noting that the decomposition we choose may be dependent on E itself, we only require that the lower bound constant does not depend on this choice of E .

1.2. A proof of the $n^{-1/2}$ bound using needle decompositions

“Needle decomposition” refers to the process of performing the step we used to obtain (1.1) until the bodies K_ω become one-dimensional. We repeatedly split the bodies in a way that the quantity $\mu_\omega(E)$ remains constant at $1/2$. Suppose that we do so and the final limiting set of needles is $(K_\omega)_{\omega \in \Omega}$. Then, we can use one-dimensional isoperimetry to get that for any ω , $\psi_\omega \gtrsim \|A_\omega\|_{\text{op}}^{-1/2}$. We also have that $\mu_\omega(E) = 1/2$, so

$$\mu^+(\partial E) \gtrsim \int_{\Omega} \|A_\omega\|_{\text{op}}^{-1/2} d\nu(\omega). \quad (1.4)$$

We wish to bound the integral on the right below.

To do so, consider (1.2) (or rather, the similar expression for the density p). Then, we have that

$$\int_{\mathbb{R}^n} p(x) x x^\top dx = \int_{\Omega} \int_{\mathbb{R}^n} p_\omega(x) x x^\top dx d\nu(\omega).$$

Thus,

$$A + b b^\top = \int_{\Omega} A_\omega + b_\omega b_\omega^\top d\nu(\omega), \quad (1.5)$$

where A and b (resp. A_ω and b_ω) refer to the covariance matrix and barycenter of μ (resp. μ_ω) respectively. Assume without loss of generality that $b = 0$. Taking the trace on either side of the above expression,

$$\begin{aligned} \text{Tr}(A) &= \int_{\Omega} \text{Tr}(A_\omega) + \|b_\omega\|^2 d\nu(\omega) \\ &\geq \int_{\Omega} \|A_\omega\|_{\text{op}} d\nu(\omega), \end{aligned}$$

where the inequality follows from the fact that A_ω is a covariance matrix so is positive semi-definite. One can then use Hölder’s inequality to get

$$\left(\int_{\Omega} \|A_\omega\|_{\text{op}} d\nu(\omega) \right) \left(\int_{\Omega} \|A_\omega\|_{\text{op}}^{-1/2} d\nu(\omega) \right)^2 \geq 1$$

and so,

$$\int_{\Omega} \|A_\omega\|_{\text{op}}^{-1/2} d\nu(\omega) \gtrsim \text{Tr}(A)^{-1/2}.$$

Substituting this back in (1.4), we get $\psi_p \gtrsim \text{Tr}(A)^{-1/2}$, that is, $\psi_n \gtrsim n^{-1/2}$.

1.3. An alternate way to look at stochastic localization

Let us return to (1.3). In the above method of needle decomposition, we attempted to exercise control over the quantity $\mu_\omega(E)(1 - \mu_\omega(E))$ for all ω by fixing $\mu_\omega(E)$ at $1/2$.

How does stochastic localization fit into this? Instead of controlling $\mu_\omega(E)$, we try to control ψ_ω by defining a martingale (p_t) whose isoperimetric constant is easily bounded. That is, $\mathbf{E}[p_t] = p$ (this is just an integral of the form of (1.1)) and further, the isoperimetric constant of μ_t is lower bounded by $t^{1/2}$. Then, the problem comes down to estimating

$$\int_{\Omega} \mu_t(E)(1 - \mu_t(E)) d\nu(\omega),$$

which is exactly what papers such as [Che21] do.

1.4. What next?

Going back to needle decompositions again, we wish to show that there exists a needle decomposition conserving $\mu_\omega(E) = 1/2$ such that

$$\int_{\Omega} \frac{1}{\|A_\omega\|_{\text{op}}^{1/2}} d\nu(\omega) \gtrsim \|A\|_{\text{op}}^{-1/2}.$$

(1.5) for $b = 0$ gives

$$\|A\|_{\text{op}} = \left\| \int_{\Omega} A_\omega + b_\omega b_\omega^\top d\nu(\omega) \right\|_{\text{op}}.$$

Therefore, it would suffice to show that

$$\int_{\Omega} \frac{1}{\|A_\omega\|_{\text{op}}^{1/2}} d\nu(\omega) \gtrsim \left\| \int_{\Omega} A_\omega + b_\omega b_\omega^\top d\nu(\omega) \right\|_{\text{op}}^{-1/2}$$

for some needle decomposition that conserves $\mu_\omega(E)$.¹

Using Hölder's inequality as we did in the proof of the $n^{-1/2}$ bound, it is seen that it suffices to show

$$\int_{\Omega} \|A_\omega\|_{\text{op}} d\nu(\omega) \lesssim \left\| \int_{\Omega} A_\omega + b_\omega b_\omega^\top d\nu(\omega) \right\|_{\text{op}}$$

for some needle decomposition preserving $\mu_\omega(E)$ (it would in fact be enough to show this with some set $A \subseteq \Omega$ instead of Ω such that $\nu(A)$ is lower-bounded by a constant).

Neglecting the $b_\omega b_\omega^\top$ term, it suffices to show that

$$\int_{\Omega} \|A_\omega\|_{\text{op}} d\nu(\omega) \lesssim \left\| \int_{\Omega} A_\omega d\nu(\omega) \right\|_{\text{op}}. \quad (1.6)$$

The above inequality essentially asks if there exists a needle decomposition where the needles are “nearly aligned”. Indeed, if the segments of the needles are perfectly aligned, then equality holds above. We are allowing a constant factor of leeway. If the direction of the one-dimensional body K_ω is u_ω , then the above is equivalent to

$$\sup_{\|\zeta\| \leq 1} \int_{\Omega} \mathbf{Var}_{x \sim p_\omega}(x) \langle \zeta, u_\omega \rangle^2 d\nu(\omega) \gtrsim \int_{\Omega} \mathbf{Var}_{x \sim p_\omega}(x) d\nu(\omega) \quad (1.7)$$

¹Is this inequality equivalent to the KLS Conjecture? Do there exist needle decompositions not obtained by the bisection method that conserve $\mu_\omega(E)$ and satisfy the above inequality?

§2. More on decompositions

2.1. Hyperplane bisections

As before, suppose we have a log-concave probability measure μ with density p on the body K , and we fix some $E \subseteq K$ with $\mu(E) = 1/2$. Let us define the function $f_{E,K} : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ by

$$f_{E,K}(x) = \left| \int_{\{z \in \mathbb{R}^n : \langle z, x \rangle \geq \|x\|\}} p(y)(\mathbb{1}_E - \mathbb{1}_{\mathbb{R}^n \setminus E}) dy \right|.$$

That is, if H_x is the hyperplane defined by x (orthogonal to x and passing through it) and H_x^+ is either of the resulting halfspaces, the value of the above function at x is equal to $|\mu(E \cap H_x^+) - \mu((\mathbb{R}^n \setminus E) \cap H_x^+)|$.

This serves as a measure of how “imbalanced” the hyperplane corresponding to x is – $f_{E,K}(x) = 0$ iff the hyperplane corresponds to x is a bisecting hyperplane (where bisecting means that $\mu(E \cap K_\omega) = \frac{1}{2}\mu(K_\omega)$, as in needle decompositions).

For nice(?) E , $f_{E,K}$ is continuous.

The primary tool used in [LS93] to prove the localization lemma was that there exists a bisecting hyperplane passing through any $(n-2)$ -dimensional affine space. How would this translate in terms of the above defined function?

Suppose we have an $(n-2)$ -dimensional affine space orthogonal to the subspace spanned by $y, z \in \mathbb{R}^n$ and passing through y .

Suppose that x defines a hyperplane containing this affine space. Then x is orthogonal to the plane, and so orthogonal to the space itself. That is, it must lie in the subspace spanned by y, z . Further, $y - x$ is orthogonal to x . That is, the set of all these x forms a circle passing through 0 contained in the 2-dimensional subspace spanned by y, z .

The conclusion of the localization method is that for any circle S passing through 0, either

- $f(w) = 0$ for some $w \in S \setminus \{0\}$ or
- The limit of $f(w)$ as w goes to 0 along the circle is equal to 0 – this corresponds to a bisecting hyperplane passing through the origin itself. It is not too difficult to check that this is well-defined and that the directional limit along either direction of the circle is the same.

More generally, suppose we have some smooth curve C in \mathbb{R}^n that passes through the origin. Then, as before, either $f(w) = 0$ for some $w \in C \setminus \{0\}$ or one of the directional limits at 0 (along C) is equal to 0.

A consequence of this nice property is that given any non-zero $x \in \mathbb{R}^n$, there exists a “barrier” curve C in \mathbb{R}^n separating x from 0 such that $f(y) = 0$ for any $y \in C$.

An interesting question is to generally characterize these functions.

2.2. Aligned 2-dimensional decompositions are always possible

Suppose we have an n -dimensional body K with $n > 2$ along with some direction u in \mathbb{R}^n . We claim that it is possible to decompose this into a set of $(n-1)$ -dimensional bodies $\{K_\omega\}$ such that any of these bodies contains our specified direction u (meaning that a translational shift of $\text{span}(\{u\})$ is contained in the minimal affine space containing any K_ω).

To prove this, assume without loss of generality that $u = e_n$. Consider the set of $(n-2)$ -dimensional affine spaces

$$S = \{\{x \in \mathbb{R}^n : x_i = q_1, x_j = q_2\} : q_1, q_2 \in \mathbb{Q}, 1 \leq i < j \leq n-1\}.$$

This is similar to the argument involved in [LS93] except that we only consider the set of $(n-2)$ -dimensional affine spaces that contain u . As the argument goes there, all the bodies must decompose into at most $(n-1)$ -dimensional bodies in the limiting step – if not, then there exists some affine space in S that intersects the n -dimensional body, and choosing the corresponding bisecting hyperplane results in a contradiction.

In fact, it turns out that we can decompose it into a set of 2-dimensional bodies that all contain our specified direction!

This is easily done using induction on n . Reducing the n -dimensional body to a set of $(n - 1)$ -dimensional bodies and then each of these smaller bodies to 2-dimensional bodies gets the job done. It should be noted that this argument does not work out if the body under consideration is 2-dimensional, since it does not make sense to have a 0-dimensional affine space containing our direction.

A natural next question is: can we give up perfect alignedness in exchange for near alignedness, which is all we really need to show KLS?

2.3. A potential function

Let us fix μ , p , K , and E as usual. Also suppose we have some direction u . We wish to decompose the body into needles in a way that all of them are nearly in the direction of u . Equivalently, the hyperplanes chosen for bisection should all nearly contain u . That is, the set of x corresponding to the hyperplanes $\{H_x\}$ must all be nearly orthogonal to u . So, at each step, the x chosen must be such that $\langle x, u \rangle$ is small – more precisely, $1 - \frac{\langle x, u \rangle^2}{\|x\|^2} \gtrsim 1$.

Also, as seen from Equation (1.3), all we really want is that $\mu_\omega(E)(1 - \mu_\omega(E)) \gtrsim 1$, it might be fine to instead just minimize $f_{E,K}$ instead of ensuring that it is exactly equal to 0. So, one may choose the x corresponding to the bisecting hyperplane at each step by constructing a potential function such as

$$\Phi(x) = (1 + f_{E,K}(x)) \left(1 + \frac{|\langle x, u \rangle|}{\|x\|} \right)$$

and at each step, choosing the x that minimizes Φ . The reason for adding the 1 is that otherwise, the expression would trivially be minimized if the corresponding term is 0 irrespective of the other term.

References

- [Che21] Yuansi Chen. An almost constant lower bound of the isoperimetric coefficient in the KLS conjecture, 2021.
- [KLS95] R. Kannan, L. Lovász, and M. Simonovits. Isoperimetric problems for convex bodies and a localization lemma. *Discrete & Computational Geometry*, 13(3):541–559, Jun 1995.
- [Led04] M. Ledoux. Spectral gap, logarithmic sobolev constant, and geometric bounds. *Surveys in differential geometry*, 9:219–240, 2004.
- [LS93] L. Lovász and M. Simonovits. Random walks in a convex body and an improved volume algorithm. *Random Structures & Algorithms*, 4(4):359–412, 1993.
- [LV18] Yin Tat Lee and Santosh S. Vempala. The Kannan-Lovász-Simonovits Conjecture, 2018.