# Day 4 Talk 3

Santosh Vempala

"Progress on the KLS Conjecture"

with slides

Joint w Yin Tat Lee

An exercise:

Lemma: isotropic log concave p:

$$\mathbb{E}_{x,y\sim p} \langle x,y\rangle^3 \lesssim n^{1.5}$$

Proof.

$$\mathbb{E}_x \mathbb{E}_y |\langle x,y\rangle|^3 \leq \mathbb{E}_x \left(\mathbb{E}_y \langle x,y\rangle^2\right)^{3/2}$$

$$= \mathbb{E}_{x\sim p} \|x\|^3$$

$$\leq \dots \quad \text{in slides}$$

---

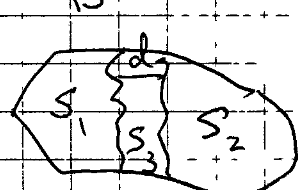
Isoperimetry, definition of Isoperimetric constant.

$$\psi_p = \min_{S: p(S)\leq \frac{1}{2}} \frac{p(\partial S)}{p(S)}$$

Conjecture: halfspaces minimizers for isoperimetric constants for log-concave densities.

(LS, DF)

~~KLS~~ thm: $p(S_3) \geq \frac{2 \cdot d}{D} \min(p(S_1), p(S_2))$



Thm KLS'95

$$p(S_3) \geq \frac{cd}{R} \min p(S_1), p(S_2)$$

---

KLS Thm & Conjecture

Conj $\psi_p \geq \frac{c}{\sqrt{\lambda_1(A)}}$

---

Outline of talk:
- Connections to other problems
- techniques
- best bounds
- open problems

---

Related to:
Algorithmic problem: Sampling

related
- min convex function
- compute volume

- complexity bounds for ball-walk depend on isoperimetric constant

- slicing conjecture

- Thin-shell conjecture

---

Bounds for $\psi_p$ imply bounds in many other problems:
- concentration
- Poincaré
- Thin-shell ...

KLS implies: $L_p \lesssim G_p \lesssim \frac{1}{\sqrt{J_p}} \approx \frac{1}{\psi_p}$

(see slide for details)

---

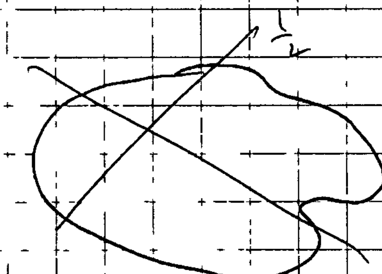
applications to metabolic networks

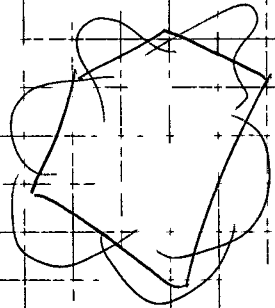
---

Techniques

- Localization & stochastic localization

Pictures from talk



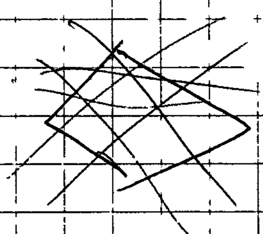



stochastic localization:

see video from slides,

$\rightsquigarrow$ gaussian

---



# The Kannan-Lovász-Simonovits Conjecture

Yin Tat Lee (U. Washington) and Santosh S. Vempala (Georgia Tech)

# Thank You!

# An exercise

Lemma. For isotropic logconcave $p$: $\mathbb{E}_{x,y\sim p}\langle x,y\rangle^3 \lesssim n^{1.5}$

Proof.

$$
\begin{aligned}
\mathbb{E}_{x\sim p}\mathbb{E}_{y\sim p}|\langle x,y\rangle|^3 &\lesssim \mathbb{E}_{x\sim p}\left(\mathbb{E}_{y\sim p}\langle x,y\rangle^2\right)^{3/2}\\
&= \mathbb{E}_{x\sim p}\|x\|^3\\
&\lesssim \left(\mathbb{E}_{x\sim p}\|x\|^2\right)^{3/2}\\
&= n^{1.5}.
\end{aligned}
$$

Exercise. Prove a better bound.

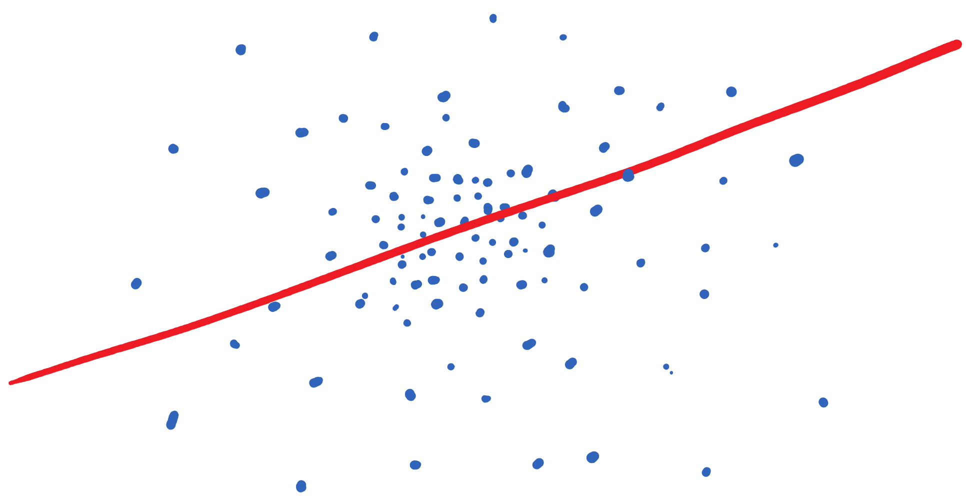
# Isoperimetry

Isoperimetric Ratio/Cheeger Constant/Expansion of a function p:

$$\psi_p = \min_{S:p(S)\leq\frac{1}{2}} \frac{p(\partial S)}{p(S)}$$

Q. What is the Cheeger constant of the Gaussian distribution?

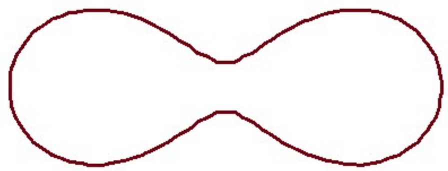
A. The isoperimetric ratio of a halfspace through its centroid: $\sqrt{\frac{2}{\pi}}$



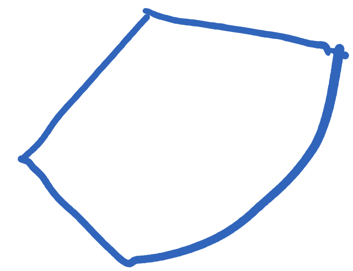
In fact for any $0 < t < 1$, the subset of measure $t$ with minimum surface area is a halfspace!

# Isoperimetry

Isoperimetric Ratio/Cheeger Constant/Expansion:



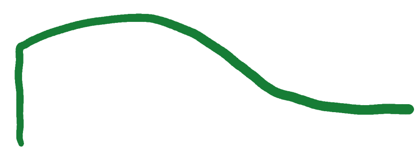
$$\psi_p = \min_{S:p(S)\leq\frac{1}{2}} \frac{p(\partial S)}{p(S)}$$



Can be arbitrarily small…
Structured distributions?

Logconcave function: $f(\lambda x + (1-\lambda)y) \geq f(x)^{\lambda} f(y)^{1-\lambda}$
(nonnegative function whose logarithm is concave)

Common generalization of Gaussians and indicators of convex sets.



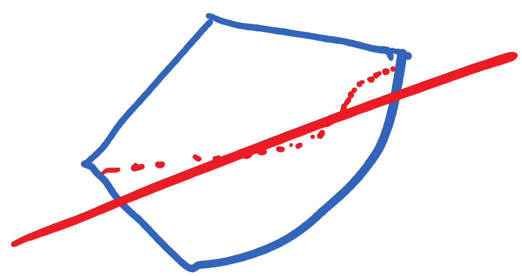
Halfspace cuts do not have to be the minimal ones, but…

# The Conjecture

Isoperimetric Ratio/Cheeger Constant/Expansion:

$$\psi_p = \min_{S: p(S) \leq \frac{1}{2}} \frac{p(\partial S)}{p(S)}$$

Logconcave function: $f(\lambda x + (1-\lambda)y) \geq f(x)^{\lambda} f(y)^{1-\lambda}$

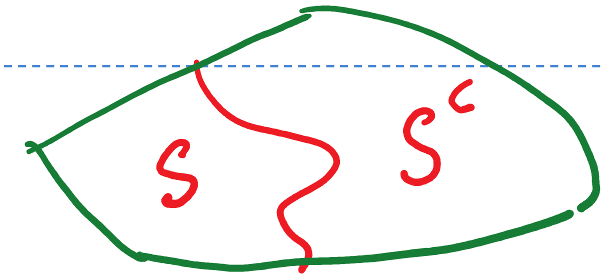


Conjecture: For any logconcave density in any dimension, halfspaces minimize the isoperimetric ratio up to an absolute universal constant.

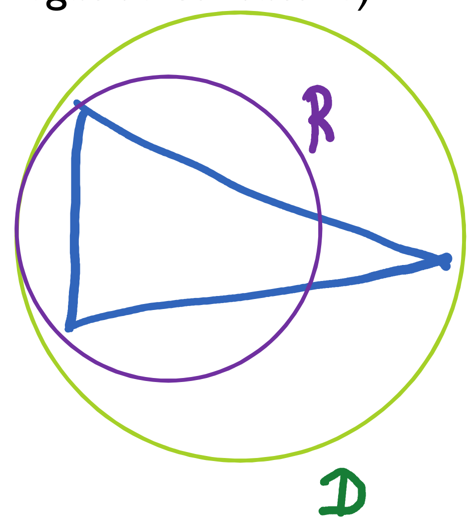
YTL1

# KLS Theorem

Thm [LS, DF] $p(\partial S) \geq \frac{2}{D}\min(p(S), p(S^c))$




(special case of isoperimetry for Riemannian manifolds with nonnegative curvature)




$A = \mathbb{E}_p((x - \bar{x})(x - \bar{x})^T)$ : covariance matrix of $p$

$R^2 = \mathbb{E}_p\left(\|x - \bar{x}\|^2\right) = Tr(A) = \sum_i \lambda_i(A)$

Thm. [KLS95]. For any logconcave density,

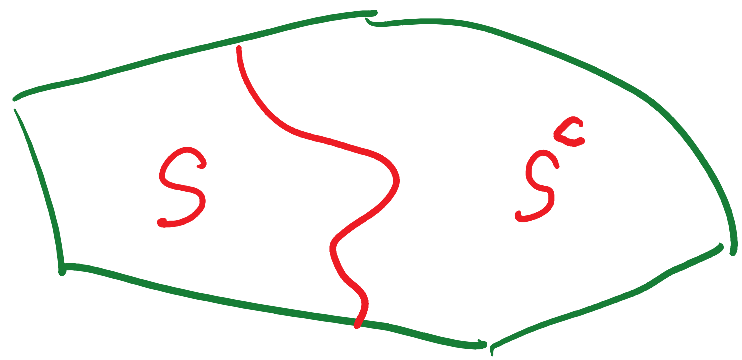
$$p(\partial S) \geq \frac{c}{R}\min(p(S), p(S^c)).$$

(note: isotropic distribution has $A = I$. So, $\psi_p \geq \frac{c}{\sqrt{n}}$ for isotropic p)

## Slide 7

**YTL1** Draw D and $R^2$?
Yin Tat Lee, 11/14/2017

# KLS Theorem and Conjecture




$A$ = covariance matrix of $p$

$$R^2 = \mathbb{E}_p\left(\|x-\bar{x}\|^2\right) = Tr(A) = \sum_i \lambda_i(A)$$

Thm. [KLS95]. $\psi_p \geq \frac{c}{\sqrt{Tr(A)}} = \frac{c}{\sqrt{n}}$ for isotropic $p$.

Conj. [KLS95]. $\psi_p \geq \frac{c}{\sqrt{\lambda_1(A)}} = \Omega(1)$ for isotropic $p$.

# Outline

Original Motivation

Connections

- Probability and Geometry
- Algorithms

Techniques

- Localization
- Stochastic Localization

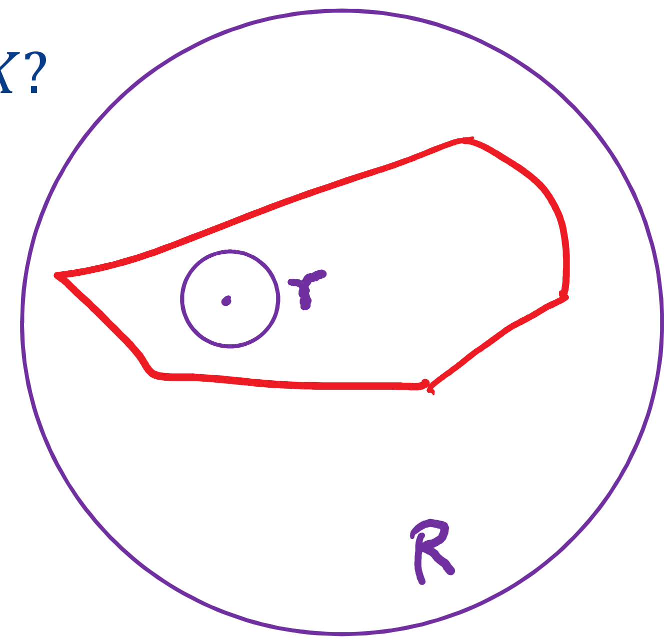
Some open problems

# An algorithmic problem: Sampling

Given convex body K, generate uniform random point in K.

K specified by a "well-guaranteed" membership oracle:

- $x_0, r, R: x_0 + rB_n \subseteq K \subseteq RB_n$
- An oracle that answers YES/NO to $x \in K$?



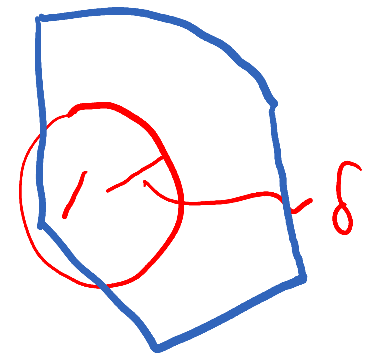

Related problems we will see later:

- Compute the volume of K
- Minimize a convex function over K

# Sampling with the ball walk

At $x$, pick random $y$ from $x + \delta B_n$,

if $y$ is in $K$, go to $y$.




Approaches the uniform distribution over K.

Rate of convergence?

Cheeger constant of Markov chain…

$$\psi_K = \min_S \frac{vol(\partial S)}{\min(vol(S), vol(S^c))}$$

Thm. [KLS97]

Mixing time of the ball walk from a warm start is $O^*\left(\frac{n^2}{\psi_K^2}\right)$.

# Connections I: Geometry and Probability

The KLS conjecture has very interesting consequences, many of which were conjectured independently and earlier.

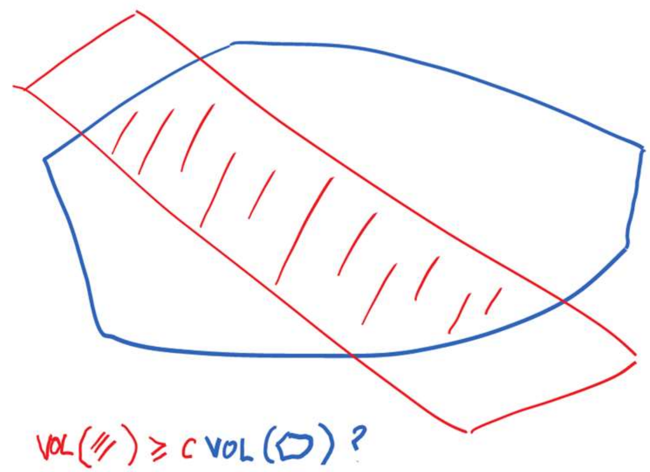
Slicing/small ball probability

Thin shell/Central Limit theorem

Poincáre/Lipschitz Concentration

# The Slicing conjecture: anti-concentration

Convex body of volume one has a hyperplane section of volume at least some constant.




Equivalently: for any isotropic logconcave density $f$,

$$L_p = f(0)^{1/n} = O(1).$$

(Paouris) Slicing implies $\Pr\left(\|x\|_2 \le \epsilon\sqrt{n}\right) \le (C\epsilon)^n$

(Ball) KLS => Slicing

(Klartag; Bourgain) $L_p \lesssim n^{1/4}$

YTL2

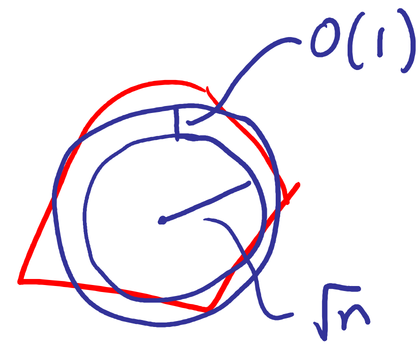
# The Thin-shell conjecture: a CLT

For X from any isotropic logconcave distribution $p$,

$$Var_p\left(\left||X|\right|^2\right) = O(n)$$

or

$$\sigma_p^2 = \mathbb{E}_p\left(\left(\left||X|\right| - \sqrt{n}\right)^2\right) = O(1)$$




"Most of an isotropic logconcave distribution is contained in an annulus of constant thickness."

KLS => thin-shell. In fact, $\sigma_p \lesssim \frac{1}{\psi_p}$

Thm [Eldan-Klartag]. $L_p \lesssim \sigma_p$

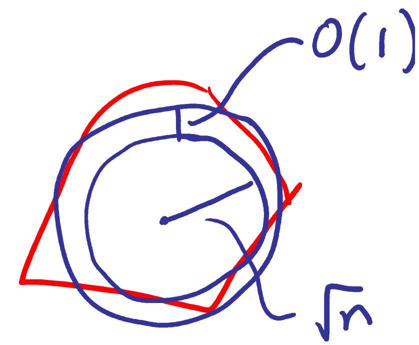
## Slide 14

**YTL2** You never defined sigma_p and L_p

Yin Tat Lee, 11/14/2017

# The Thin-shell conjecture: a CLT

X from any isotropic logconcave distribution,




$$\mathbb{E}\left(\left(\left|\left|X\right|\right| - \sqrt{n}\right)^2\right) = O(1)$$

$$\sigma_p \lesssim \frac{1}{\psi_p}$$

[Eldan-Klartag]. $L_p \lesssim \sup_p \sigma_p$

CLT: Most marginals are approximately Gaussian.

For an isotropic convex body K, let $g_\theta(s) = vol\left(K \cap \left\{x: x^T\theta = s\right\}\right)$
Then,

$$\Pr\left(\left\{\theta \in S^{n-1}: \max_t \left|\int_{-\infty}^{t} g_\theta(s) - \int_{-\infty}^{t} \gamma(s)\right| \lesssim \delta + \frac{\psi_K}{\sqrt{n}}\right\}\right) \geq 1 - n\exp\left(-\Omega\left(\delta^2 n\right)\right)$$

# Progress on the thin-shell bound

| Year/Authors | bound on $\sigma_p$ |
| --- | --- |
| 2006/Klartag | $\sqrt{\dfrac{n}{\log n}}$ |
| 2006/Fleury-Guédon-Paouris | $\sqrt{n}\,\dfrac{(\log\log n)^2}{\log^{1/6} n}$ |
| 2006/Klartag | $n^{2/5}$ |
| 2010/Fleury | $n^{3/8}$ |
| 2011/Guédon-E. Milman | $n^{1/3}$ |
| 2016/Lee-Vempala | $n^{1/4}$ |

# Poincáre Conjecture

Smooth function g

Isotropic logconcave density p

There exists a universal constant c s.t.

$$\forall g: \ \ \zeta_p = \inf \frac{\mathbb{E}_p\left(\|\nabla g\|^2\right)}{Var_p(g)} \geq c$$

Thm. [Mazja,Cheeger; Buser; Ledoux] $\quad \zeta_p \approx \psi_p^2$

# Lipschitz concentration

Lipschitz function g on the sphere

Levy's classical concentration:

$$\Pr\left(\left|g(x) - \mathbb{E}_{S^{n-1}}(g)\right| \geq t\right) \leq 2e^{-ct^2 n}$$

Thm.[Gromov-Milman] L-Lipschitz function g in $R^n$,
isotropic logconcave density p,

$$\Pr_p(|g(x) - \mathbb{E}(g)| > L \cdot t) \leq e^{-\Omega(t\psi_p)}$$

Thm. [E. Milman] Concentration ⇒ KLS.

# Entropy gaps and jumps

$\mathrm{Ent}(X) = -E(\log p(x))$. How random is a distribution?

[Shannon-Stam] Entropy gap: equality only for Gaussian

$$X, Y \sim p: \qquad \mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) \geq \mathrm{Ent}(X)$$

[Ball, Nguyen] Entropy gap jump. $X, Y \sim p,\ Z \sim N(0, I)$

$$\mathrm{Ent}\left(\frac{X+Y}{\sqrt{2}}\right) - \mathrm{Ent}(X) \gtrsim \psi_p^2\big(\mathrm{Ent}(Z) - \mathrm{Ent}(X)\big)$$

# Connections: Geometry and Probability

Slicing conjecture: $L_p = p(0)^{1/n} = O(1)$

Thin-Shell conjecture: $\sigma_p = \mathbb{E}\left(\|x\| - \sqrt{n}\right)^2 = O(1)$

Poincáre conjecture: $\zeta_p = \inf_g \frac{\mathbb{E}_p\left(\|\nabla g\|^2\right)}{Var_p(g)} = \Omega(1)$

Generalized Levy concentration:
Lipschitz $f$ with $\mathbb{E}f = 0,\ \mathbb{P}(f(x) > t) = \exp(-\Omega(t))$.

KLS conjecture implies all: $L_p \lesssim \sigma_p \lesssim \frac{1}{\sqrt{\zeta_p}} \approx \frac{1}{\psi_p}$

# Connections: Algorithms

Sampling

Optimization

Volume Computation/Integration

Learning

What is the complexity of computational problems as the dimension grows?

Dimension = number of variables

Typically, size of input is a function of the dimension.

# Computational model

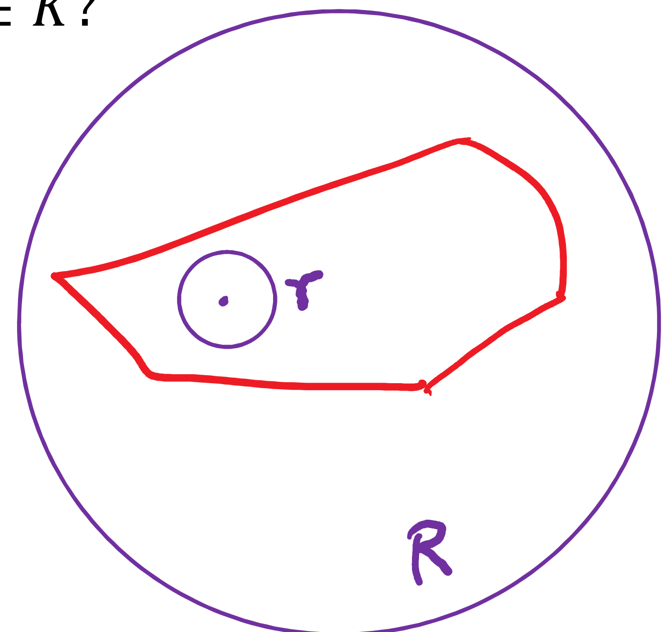
## Well-guaranteed Membership oracle:

Compact set K is given by

a membership oracle: answers YES/NO to "$x \in K$?"

a point $x_0 \in K$

Numbers r, R s.t. $x_0 + rB^n \subseteq K \subseteq RB^n$




## Well-guaranteed Function oracle

An oracle that returns $f(x)$ for any $x \in R^n$

A point $x_0$ with $f(x_0) \geq \beta$

Numbers r, R s.t.

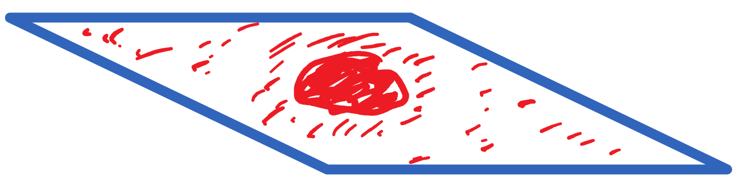
$$x_0 + rB^n \subset L_f\left(\frac{1}{8}\right) \text{ and } R^2 = \mathbb{E}_f\left(\left\|X - \bar{X}\right\|^2\right)$$

# Problem 1: Sampling

Input: function f: $R^n \rightarrow R_+$, $\int f < \infty$, specified by an oracle, a point x, error parameter ε.

Output: A point y from a distribution within distance ε of distribution with density proportional to f.

Examples: $f(x) = 1_K(x)$, $f(x) = e^{-a||x||}1_K(x)$

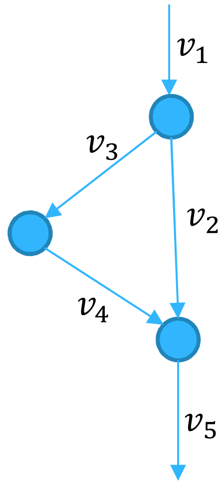


# Sampling metabolic networks

Given a metabolic network $S \in \mathbb{R}^{m \times n}$ on $m$ metabolites and $n$ reactions.

Find mass conserving flow $\mathrm{v} \in \mathbb{R}^n$ with bounds $l, u \in \mathbb{R}^n$:

$$Sv = b$$
$$l \leq v \leq u$$




Many possible $v$'s, which one to pick?

- Sample over all possible values of $v$!

Dimension: 5000-100,000.

# Analysis of metabolic networks

Sampling enables an unbiased study of all feasible metabolic flows

- Could optimize with respect to an objective function, but it is unclear if the human body acts like this

Can also compute the volume of the space!

- Price et al. (2004) analyze the human red blood cell metabolic network (dim=11)
- They observe the volume of diseased patient's networks is significantly lower.

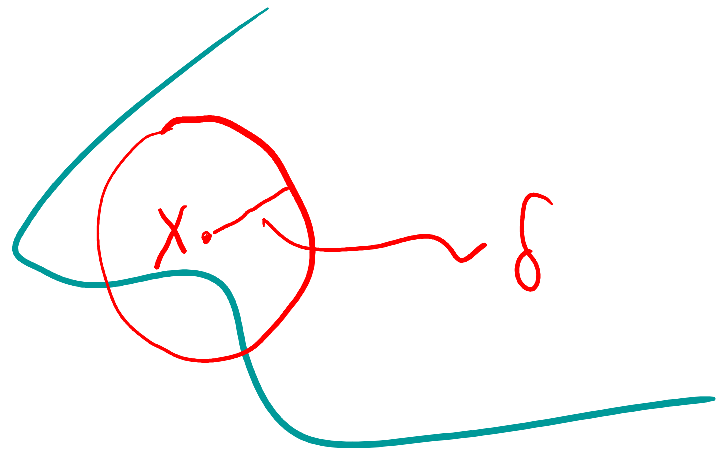
# How to Sample?

Ball walk:

At x,

-pick random y from $x + \delta B_n$

-if y is in K, go to y

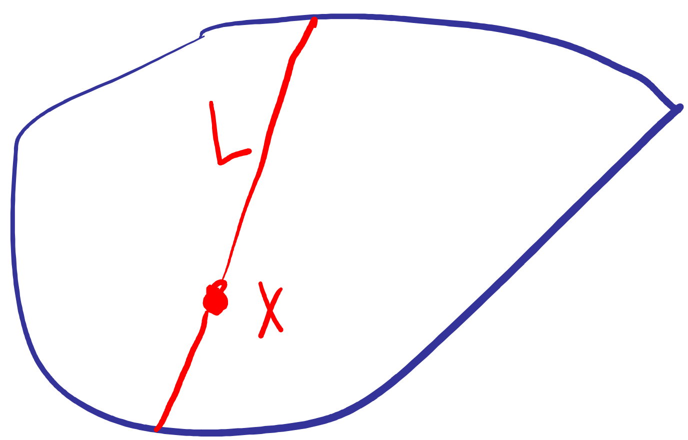



Hit-and-Run:

At x,

-pick a random chord L through x

-go to a random point y on L



# Markov chains

State space K , next step distribution $P_u(.)$ associated with each point u in K.
Stationary distribution Q, ergodic "flow" defined as

$$\Phi(A) = \int_A P_u(K\backslash A) dQ(u)$$

For a stationary distribution, we have $\Phi(A) = \Phi(K\backslash A)$
Conductance:

$$\phi(A) = \frac{\int_A P_u(K\backslash A) dQ(u)}{\min Q(A),\ Q(K\backslash A)} \qquad \phi = \inf\ \phi(A)$$

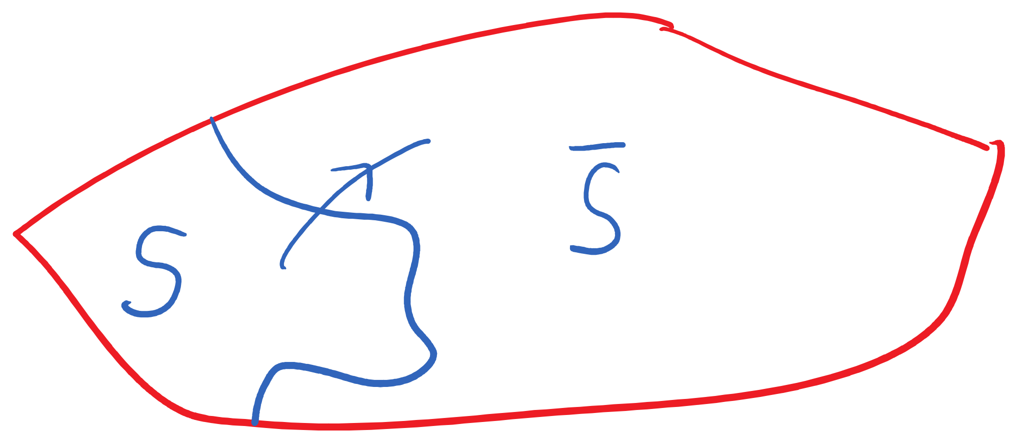
Thm. [LS93] $Q_t$: distribution after t steps

$$M = \sup_{A\subset K} \frac{Q_0(A)}{Q(A)}: \quad d_{TV}(Q_t, Q) \le \sqrt{M}\left(1 - \frac{\phi^2}{2}\right)^t$$

$$M = E_{Q_0}\left(\frac{Q_0(x)}{Q(x)}\right): \ d_{TV}(Q_t, Q) \le \epsilon + \sqrt{\frac{M}{\epsilon}}\left(1 - \frac{\phi^2}{2}\right)^t \ \forall \epsilon > 0$$

# Conductance

Consider an arbitrary measurable subset S.




Need to show that the escape probability from S is large.

(Smoothness of 1-step distribution) Points that do not cross over are far from each other i.e., nearby points have large overlap in 1-step distributions

(Isoperimetry) Large subsets have large boundaries

# Convergence of ball walk

Theorem [KLS97]. The ball walk applied to an isotropic logconcave density p, from a warm start, converges in $O^*\left(\frac{n^2}{\psi_p^2}\right)$ steps.

"Cheeger constant of this Markov chain is determined by Cheeger constant of its stationary distribution"

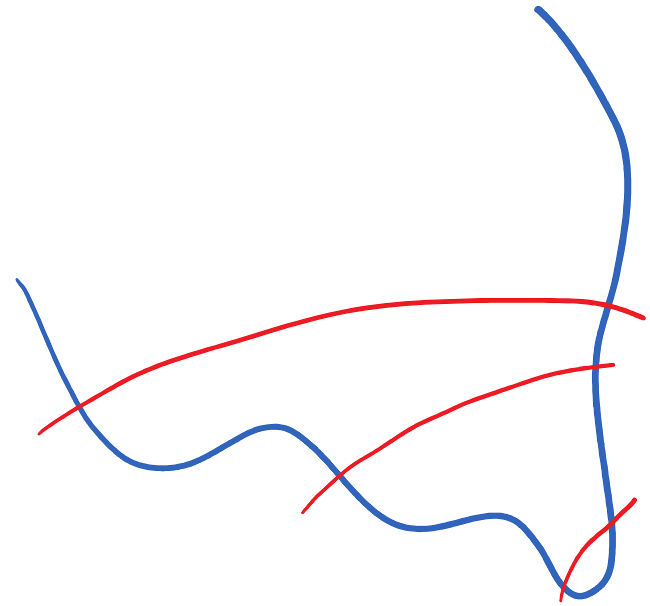
# Problem 2: Optimization

Input: function f: $R^n \rightarrow R$ specified by an oracle,

point x, error parameter ε .

Output: point y such that

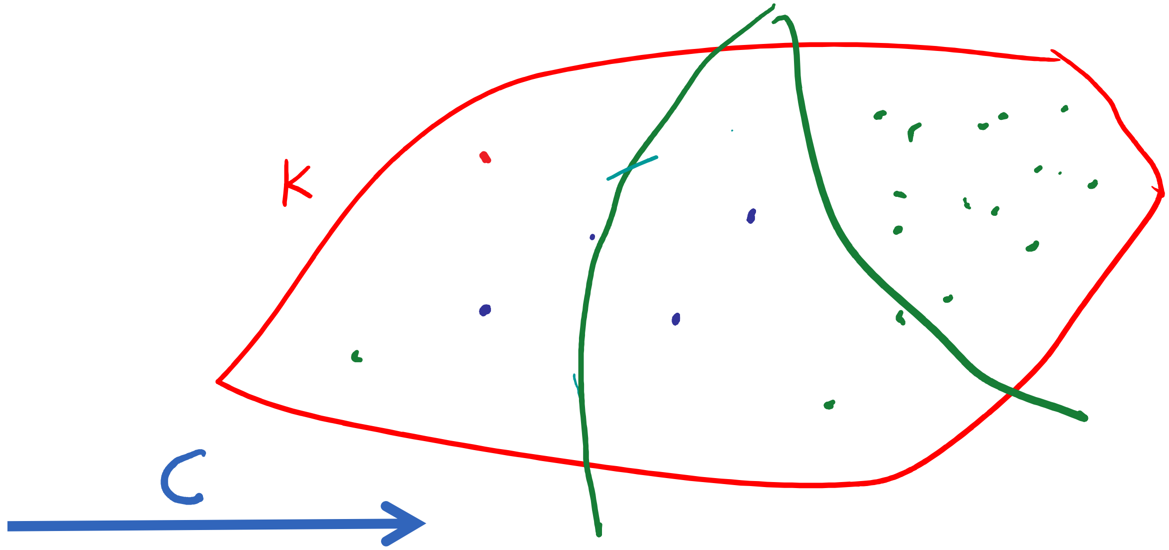
$$f(y) \geq \max f - \epsilon$$

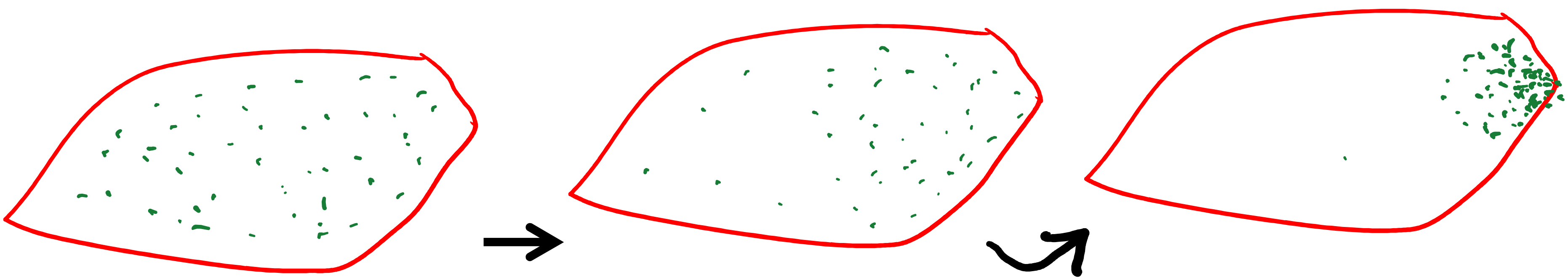


Examples: $\max c \cdot x \; s.t. Ax \geq b$ , $\min ||x|| \; s.t. x \in K$.

# Optimization from membership

Sampling suggests a conceptually very simple algorithm.



# Simulated Annealing [Kalai-V.04]

To optimize $f$ consider a sequence $f_0, f_1, f_2, \dots,$
with $f_i$ more and more concentrated near the optimum.

$$f_i(x) = e^{-t_i \langle c,x \rangle}$$

Corresponding distributions:

$$P_{t_i}(x) = \frac{e^{-t_i \langle c,x \rangle}}{\int_K e^{-t_i \langle c,x \rangle} dx}$$

Lemma. $E_{P_t}(c \cdot x) \leq \min c \cdot x + \frac{n}{t}.$

So going up to $t = \frac{n}{\epsilon}$ suffices to obtain an $\epsilon$ approximation.

# Volume Computation

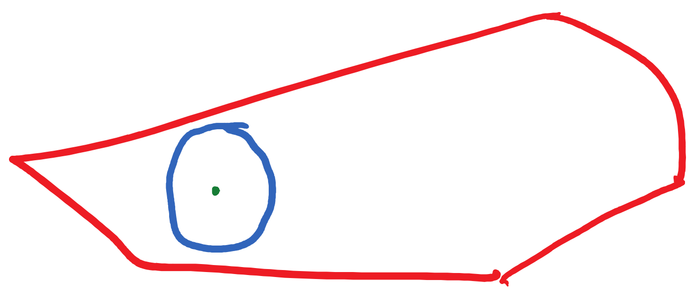
Given a measurable, compact set K in n-dimensional space and $\epsilon > 0$, find a number A such that:

$$(1-\epsilon)\,\text{volume}(K) \;\le\; A \;\le\; (1+\epsilon)\,\text{volume}(K)$$

K is given by

a point $x_0 \in K$, s.t. $x_0 + B_n \subseteq K \subseteq RB_n$

a membership oracle: answers YES/NO to "$x \in K?$"



# *Randomized* Volume/Integration

[DFK89]. Polytime randomized algorithm that estimates volume to within relative error $(1+\epsilon)$ with probability at least $1-\delta$ in time poly(n, $\frac{1}{\epsilon}$, $\log\left(\frac{1}{\delta}\right)$).

[Applegate-K91]. Polytime randomized algorithm to estimate integral of any (Lipshitz) logconcave function.

# Progress on Volume Computation

| | Power | New aspects |
|---|---|---|
| Dyer-Frieze-Kannan 89 | 23 | everything |
| Lovász-Simonovits 90 | 16 | localization |
| Applegate-K 90 | 10 | logconcave integration |
| L 90 | 10 | ball walk |
| DF 91 | 8 | error analysis |
| LS 93 | 7 | multiple improvements |
| KLS 97 | 5 | speedy walk, isotropy |
| LV 03,04 | 4 | annealing, isoperimetry |
| LV 06 | 4 | integration, local analysis |
| Cousins-V. 15 (well-rounded) | 3 | Gaussian cooling |

# Does it work?

[Cousins-V.13]  Matlab implementation of a new algorithm

- "volume computation matlab"
- https://volumecomputation.wordpress.com/
- Incorporated into the COBRA toolbox for Systems Biology

# Outline

Original Motivation

Connections

- Probability and Geometry
- Algorithms

Techniques

- Localization
- Stochastic Localization

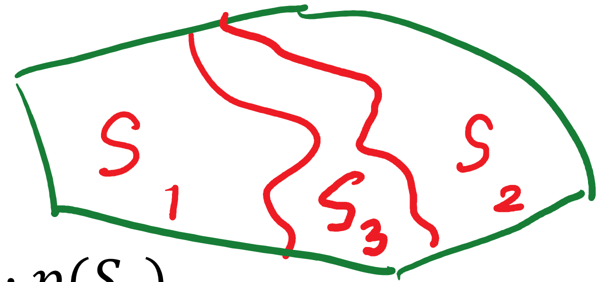
Some open problems

# Localization

Idea: Reduce inequalities in high dimension to inequalities in one dimension.

# Isoperimetry via localization

$$p(S_3) \geq \frac{2d(S_1,S_2)}{D} \min p(S_1), p(S_2)$$




Write as 2 inequalities: $p(S_1) \leq p(S_2),\ \ p(S_3) \geq \psi \cdot p(S_1)$

Let $g(x) = f(x)(1_{S_2}(x) - 1_{S_1}(x)), \quad h(x) = f(x)(\psi \cdot 1_{S_1}(x) - 1_{S_3}(x))$

Then, need to show: $\int g \geq 0 \Rightarrow \int h \leq 0.$

Suppose not, i.e., $\exists S_1, S_2, S_3$: $\int g \geq 0, \int h > 0.$

Idea:

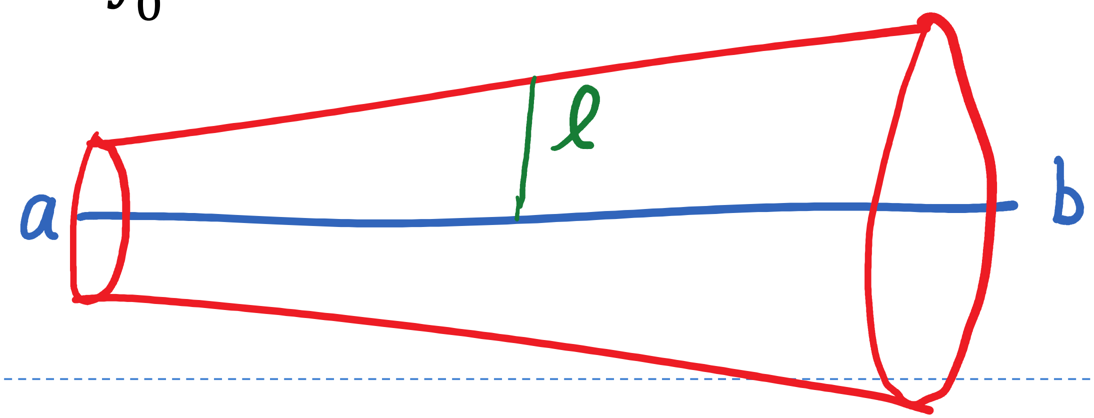
1. No such counterexample in one dimension
2. If such a counterexample exists in some dimension, then it also exists in 1 dimension.

# Localization Lemma [LS, KLS]

Lemma. Let $g, h: R^n \to R$ be integrable, lower semi-continuous functions. Suppose $\int g, \ \int h > 0$. Then, there exists an interval $[a, b] \subset R^n$ and a linear function $\ell: [0,1] \to R_+$ s.t.

$$\int_0^1 g\big((1-t)a + tb\big)\ell(t)^{n-1}\, dt > 0$$

$$\int_0^1 h\big((1-t)a + tb\big)\ell(t)^{n-1}\, dt > 0.$$



# Localization lemma

$g, h: R^n \to R, \ \ \int g, \ \int h > 0.$

1. Find a bisecting halfspace for one function
2. Show support of limit of bisections is an interval or a point.
3. The limit function has a concave profile
4. Reduce to linear cross-sectional profile.

[Fradelizi-Guédon] Extremal characterization and generalization to multiple inequalities.

# Isoperimetry via localization

$$vol(S_1) \leq vol(S_2) \Rightarrow vol(S_3) \geq \psi\, vol(S_1)$$

$$g(x) = 1_{S_2}(x) - 1_{S_1}(x), \qquad h(x) = \psi\, 1_{S_1}(x) - 1_{S_3}(x)$$

Need to show: $\int g \geq 0 \Rightarrow \int h \leq 0$.

Suppose not, i.e., for some partition, $\int g \geq 0, \int h > 0$.

Applying localization,

$$\int_0^1 g\big((1-t)a + tb\big)\ell(t)^{n-1}\, dt > 0, \int_0^1 h\big((1-t)a + tb\big)\ell(t)^{n-1}\, dt > 0.$$

Let $Z_i = \{t \in [0,1]: (1-t)a + tb \in S_i\},\ F(t) = f\big((1-t)a + tb\big)\ell(t)^{n-1}$.

Then this means that

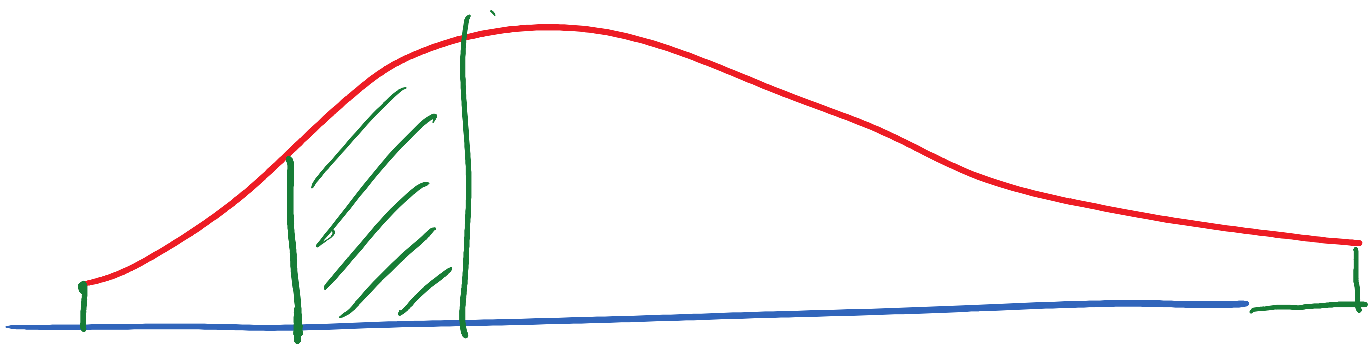
$\int_{Z_1} F \leq \int_{Z_2} F,$ but $\int_{Z_3} F < \psi \int_{Z_1} F,$ a 1-d counterexample must exist.

# One-dimensional isoperimetry

For any logconcave function:

$$\int_{S_3} F \geq \frac{2d(S_1, S_2)}{D} \min \int_{S_1} F , \int_{S_2} F$$

Suffices to show it for partition of into 3 intervals.



Without factor of 2, follows from unimodality!
Therefore, same isoperimetric ratio holds in $R^n$.

# Localization: many applications

Many other isoperimetric inequalities

E.g. the KLS conjecture holds for a Gaussian restricted by any logconcave function.

Thm. For a density h proportional to $e^{-\frac{t}{2}\|x\|^2} f(x)$ for any logconcave function f, we have $\psi_p \gtrsim \sqrt{t}$.

Analysis of [Cousins-V.2015] algorithm:

Thm. Volume of a well-rounded convex body ($B \subseteq K \subseteq \tilde{O}(\sqrt{n})$) can be computed in $O^*(n^3)$ steps!

Carbery-Wright anti-concentration of polynomials

# Stochastic localization

Goal: Lower bound on expansion of subset S of measure ½.

Idea: Apply hyperplane bisections randomly. Show measures of sets remain close to original. Prove isoperimetry for (hopefully) simpler distribution.

*(I tried this for years, still think it might work ☺)*

*(I also think there will be world peace, and I will stop eating junk food tomorrow)*

Meanwhile, Eldan: Apply infinitesimal linear reweighting in random direction to maintain expected value of density at each point.

[Eldan2012] used this to prove that the thin-shell conjecture implies the KLS constant up to a logn factor.

# Eldan's Stochastic Localization

The process starts with $p_0(x) = p(x)$ and maintains a density $p_t$ at time t with mean $\mu_t = E_{p_t}(x)$.

The infinitesimal change is

$$dp_t(x) = (x - \mu_t)^T dW_t \cdot p_t(x)$$

where $dW_t$ is an infinitesimal Gaussian (Wiener process).

We can imagine this discretely as

$$p_{t+h}(x) = p_t(x)\left(1 + \sqrt{h}(x - \mu_t)^T W_t\right)$$

where $W_t \sim N(0, I)$ is a standard Gaussian.

# Stochastic localization apps

Thm [Eldan12]. Let $\sigma(n) = \sup_p \sigma_p$. Then,

$$\psi_p \gtrsim \frac{\log n}{\sigma(n)}$$

Thm [LeeV16]. For any logconcave density p with covariance A:

$$\psi_p \gtrsim \frac{1}{Tr(A^2)^{1/4}}$$

For isotropic logconcave p, $\psi_p \gtrsim n^{-1/4}$.

Thm [LeeV17]. Log-Sobolev constant of isotropic logconcave p with support of diameter D is $[[Ӝ nexttalk Ӝ]]$. This bound is tight.

# The KLS constant

KLS theorem.

$$\psi_p \gtrsim \frac{1}{Tr(A)^{1/2}} = \frac{1}{\left(\sum_i \lambda_i(A)\right)^{1/2}}$$

LeeV theorem.

$$\psi_p \gtrsim \frac{1}{Tr(A^2)^{1/4}} = \frac{1}{\left(\sum_i \lambda_i(A)^2\right)^{1/4}}$$

KLS conjecture.

$$\psi_p \gtrsim \frac{1}{\|A\|_{op}^{1/2}} = \frac{1}{\lambda_1(A)^{1/2}}$$

# Proof Strategy

$p_t$ is a martingale:

$$\mathbb{E}p_t = p_0$$

Suffices to prove the theorem for $p_t$:

$$\psi_p = \Omega(\psi_{p_T}) \text{ for "large" } T.$$

$p_t$ has large Cheeger constant:

$$\psi_{p_T} \gtrsim \sqrt{T}$$

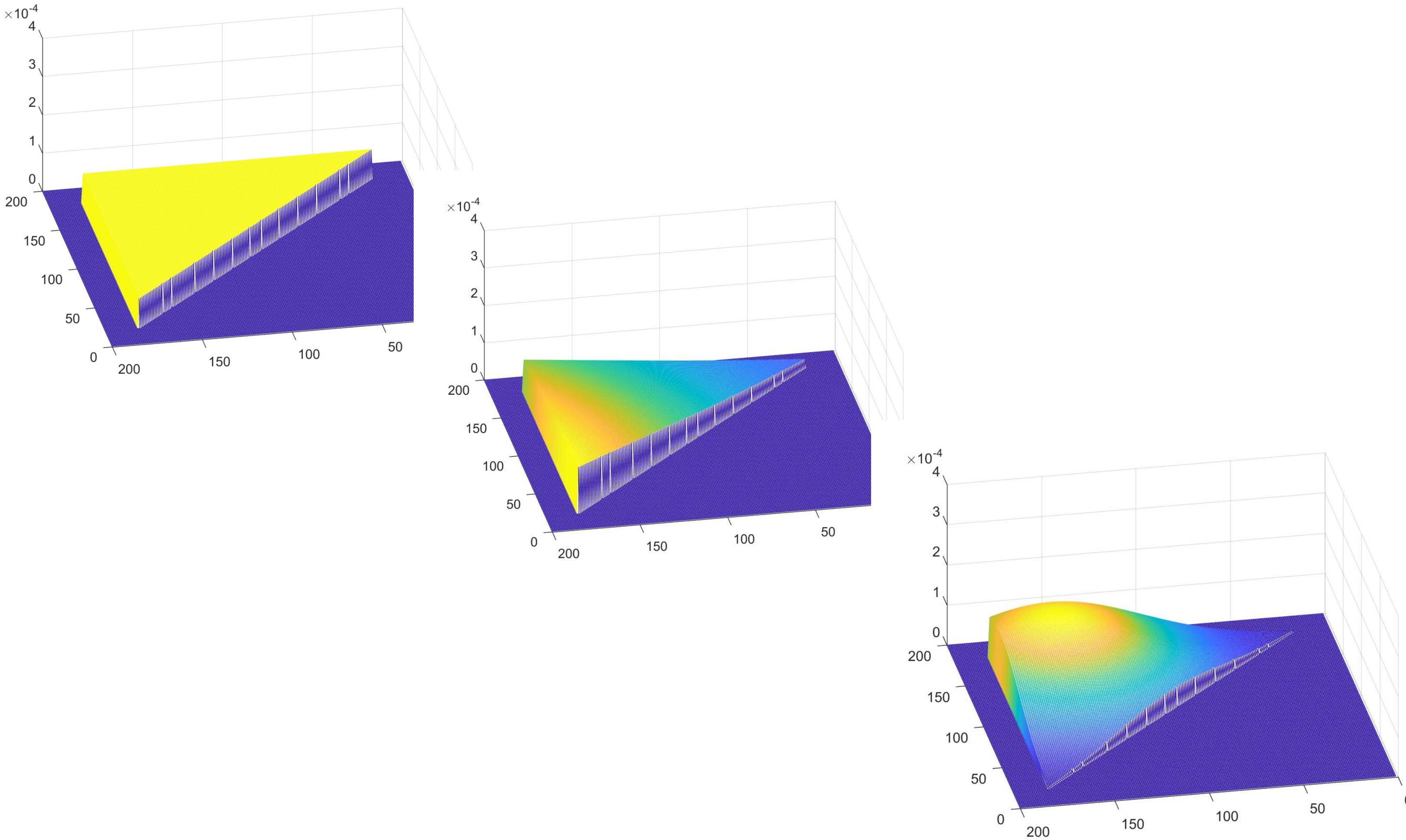
# Why does $\boldsymbol{p_T}$ have good expansion?

Localization

Stochastic localization

But's let's see it for real…

# What is happening?



# Really?

# Emergence of Gaussian factor

$$dp_t(x) = (x-\mu_t)^T dW_t \cdot p_t(x)$$

We will see that:

$$d\log p_t(x) \ = (x-\mu_t)^T dW_t - \frac{1}{2}\|x-\mu_t\|^2 dt$$

Itô's lemma: $dX_t = \mu_t dt + \sigma_t dW_t \Rightarrow \ df(X_t) = \frac{df(X_t)}{dX} dX_t + \frac{1}{2}\frac{d^2 f(X_t)}{dx^2}\sigma_t^2 dt$

This is by Taylor expansion and noting that $(dW_t)^2 = dt$.
Applying this,

$$\begin{aligned} d\log p_t(x) \ &= \frac{dp_t(x)}{p_t(x)} - \frac{1}{2}\frac{\left(\|x-\mu_t\| \cdot p_t(x)\right)^2}{p_t(x)^2} dt \\ &= (x-\mu_t)^T dW_t - \frac{1}{2}\|x-\mu_t\|^2 dt \\ &= x^T(\mu_t dt + dW_t) - \frac{1}{2}\|x\|^2 dt + g(t) \end{aligned}$$

where the last term does not depend on x.
Therefore,

$$p_t(x) \propto e^{x^T c_t - \frac{t}{2}\|x\|^2} p(x) \quad \text{for} \ \ dc_t = \mu_t dt + dW_t$$

# KLS is easy for Gaussianic distribution

Thm [Bakry-Ledoux96, also Bobkov2000, Cousins-V. 2013 by localization]

For a density h proportional to $e^{-\frac{t}{2}\|x\|^2} f(x)$ for any logconcave function f, we have $\psi_p \gtrsim \sqrt{t}$.

Proof.

Apply localization lemma.

The resulting statement in 1-d is implied by the following Brascamp-Lieb inequality: the variance of a density given by a Gaussian times a logconcave function (in one dimension) is at most the variance of the Gaussian.

Recall that $p_t(x) \propto e^{x^T c_t - \frac{t}{2}\|x\|^2} p(x)$ has such a Gaussian part!

Hence, $\psi_{p_t} = \Omega(\sqrt{t})$.

# How long can we go?

$p_t$ is a martingale: $dp_t(x) = (x - \mu_t)^T dW_t \cdot p_t(x)$

Let $A_t$ be the covariance of $p_t$.
For any measurable subset E,

$$\frac{d}{dt}\int_E p_t(x)dx = \int_E (x - \mu_t)^T dW_t p_t(x)dx = \left(\int_E (x - \mu_t) p_t(x)dx\right)^T dW_t$$

Hence, the measure of E (or any subset) is also a martingale.

$$\begin{aligned} Var\left(\frac{d}{dt}\int_E p_t(x)dx\right) &= \left\|\int_E (x - \mu_t) p_t(x)dx\right\|_2^2 \\ &\le \max_{\|\zeta\|_2 \le 1} \left(\int_E (x - \mu_t)^T \zeta \cdot p_t(x)dx\right)^2 \\ &\le \max_{\|\zeta\|_2 \le 1} \int_E \left((x - \mu_t)^T \zeta\right)^2 \cdot p_t(x)dx \\ &= \|A_t\|_{op} \end{aligned}$$

As long as $\|A_t\|_{op}$ is bounded, any set is approximately preserved.

# Suffices to bound $\boldsymbol{\psi}$ for $p_t$

$$dp_t(x) = (x - \mu_t)^T dW_t \cdot p_t(x)$$

Thm [E. Milman09] To bound $\psi$ it suffices to consider subsets of measure ½.

Suppose $\int_0^T \|A_t\|_o \; dt \leq 0.01$ with constant probability.

Since $\mathbb{E}p_T = p_0$, we have that

$$\begin{aligned}\frac{p(\partial S)}{p(S)} &= 2\mathbb{E}p_T(\partial S) \\ &\geq 2\psi_{p_T}\mathbb{E}(\min p_T(S), p_T(S^c)) \\ &\geq \frac{2\psi_{p_T}}{4}\Pr\left(\frac{1}{4} < p_T(S) < \frac{3}{4}\right) \\ &= \Omega(\psi_{p_T}) = \Omega\left(\sqrt{T}\right).\end{aligned}$$

Need to keep the spectral norm of covariance small for as long as possible…

# Bounding $\|A_t\|_{op}$

Back where we started?!

The stochastic process will give us some control.
We use the potential $\phi_t = \mathrm{Tr} A_t^2$.

We will see that $\mathrm{Tr} A_t^2 \leq 2\phi_0$ when $t \lesssim 1/\sqrt{\phi_0}$ with high probability.

Therefore, if $T \leq 0.001/\sqrt{\phi_0}$, we have $\int_0^T \|A_t\|_{op} dt \leq 0.01$.

This gives $\psi_p \gtrsim \sqrt{T} \gtrsim \left(\mathrm{Tr} A^2\right)^{-1/4}$.

- For isotropic p, this is $\psi_p \gtrsim n^{-1/4}$.

# Bounding the largest eigenvalue

$$\phi_t = \mathrm{Tr} A_t^2$$

Let's Itô it!

$$d\phi_t = 2E(x-\mu_t)^T A_t (x-\mu_t)(x-\mu_t)^T dW_t - 2\mathrm{Tr} A_t^3 dt + E((x-\mu_t)^T(y-\mu_t))^3 dt$$

$$= \delta_t dt + v_t^T dW_t$$

Lemma. For a logconcave density p,

- $\mathbb{E}\left(\|x\|^k\right) \leq (2k)^k \mathbb{E}\left(\|x\|^2\right)^{k/2}$
- $\mathbb{E}\,|(x-\mu)^T(y-\mu)|^3 \lesssim \left(\mathrm{Tr}A^2\right)^{3/2}$
- $\left\|\mathbb{E}(x-\mu)(x-\mu)^T A(x-\mu)\right\| \lesssim \|A\|_{op}^{1/2} \cdot \mathrm{Tr}A^2$

Using this, $|\delta_t| \lesssim \phi_t^{3/2}$ and $\|v_t\| \lesssim \phi_t^{5/4}$

Therefore, $d\phi_t \lesssim \phi_t^{3/2} dt + \phi_t^{5/4} dW_t$

# Bounding the largest eigenvalue

$$\phi_t = \mathrm{Tr} A_t^2$$

After Itôing,

$$d\phi_t \lesssim \phi_t^{3/2} dt + \phi_t^{5/4} dW_t$$

Or

$$\phi_t \lesssim \phi_t^{3/2} t + \phi_t^{5/4} \sqrt{t}$$

So $\phi_t \leq 2\phi_0$ for $t \lesssim 1/\sqrt{\phi_0}$

And for $T = c/\sqrt{\phi_0}$, we get:

$$\int_0^T \|A_t\|_o \ dt \leq \int_0^T \sqrt{\phi_t}\, dt \leq \sqrt{2\phi_0} \cdot T \leq 0.01.$$

So the measure of the subset stays balanced up to time $T = c/\sqrt{\phi_0}$ and the lower bound on expansion is

$$\Omega\left(\sqrt{T}\right) \gtrsim \left(\mathrm{Tr} A^2\right)^{-1/4}$$

(This is $\gtrsim n^{-1/4}$ for isotropic p.)

# An improved concentration inequality

Thm. [Paouris2006]. For an isotropic logconcave p,

$$\Pr(\|x\| - \sqrt{n} > c \cdot t) \leq e^{-t}$$

Best possible when $t \gtrsim \sqrt{n}$.

[Guédon-E. Milman]: RHS is $\exp(-\min(\frac{t^3}{n}, t))$.

[Lee-V. 2017] RHS is $\exp(-\min(\frac{t^2}{\sqrt{n}}, t))$

Thm. For any isotropic logconcave p, and any Lipschitz function g,

$$\Pr(|g(x) - \bar{g}| > c \cdot t) \leq 2e^{-\frac{t^2}{t+\sqrt{n}}}$$

# Outline

Original Motivation

Connections

- Probability and Geometry
- Algorithms

Techniques

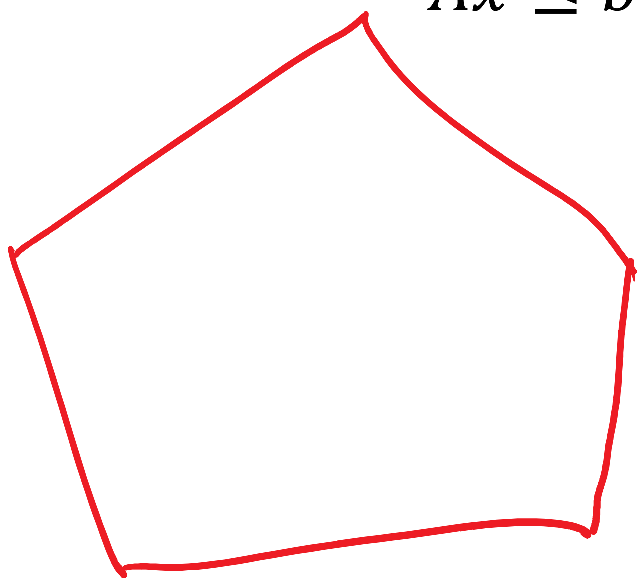
- Localization
- Stochastic Localization

Some open problems

# Deterministic Polytope Volume

Can we estimate the volume of an explicit polytope in *deterministic* polynomial time?

$$Ax \leq b$$



# Lower bound for Sampling

KLS says complexity of sampling from a warm start is $n^2$.

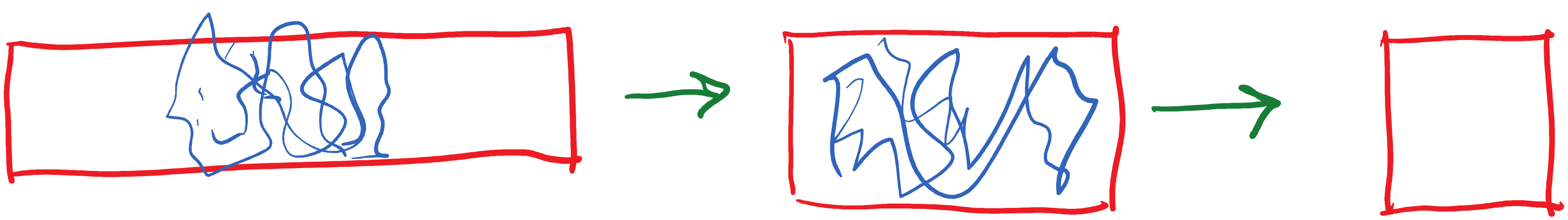
Is this the best possible?

# Faster isotropy and sampling

Isotropic transformation/rounding is the bottleneck for faster general volume computation/sampling.

Candidate algorithm:

Repeat:
1. Estimate the covariance of the standard Gaussian density restricted to the current convex body.
2. If the covariance has eigenvalues smaller than some constant, apply a transformation to make it identity.



Conjecture [Cousins-V.]. This algorithm terminates in O(log n) iterations with a well-rounded body.

# Faster isotropy and sampling

Per-step arithmetic complexity: $n^2$.

Coordinate Hit-and-Run. Could be faster by a factor of n in the per-step complexity.

But is it rapidly mixing? n^3?!

# Manifold KLS

Thm. [Lee-V. 2017] Let K be a convex body and $\phi: K \to R$ be a convex function with a convex Hessian. Let d be the distance in the Riemannian metric defined by the Hessian. Then, for any partition of K into subsets $S_1, S_2, S_3$,

$$\frac{\int_{S_3} e^{-\alpha\phi(x)} dx}{\min \int_{S_1} e^{-\alpha\phi(x)} dx \,, \int_{S_2} e^{-\alpha\phi(x)} dx} \gtrsim \sqrt{\alpha} d(S_1, S_2)$$

In other words, this Gibbs distribution satisfies a manifold KLS! $\phi(x) = \|x\|^2$ and Euclidean metric d is the special case of KLS for a Gaussian restricted to a convex body.

What are further generalizations?

# Needle decompositions

Used by [Bobkov]; also [Chandrasekaran-Dadush-V.]

Apply hyperplane cuts to get a needle decomposition

Maintain relative measure of subset S.

Show that a positive fraction of needles have bounded variance.

Conclude KLS!