Reducing Sampling to KLS

He Jia, Aditi Laddha, Yin Tat Lee, Santosh Vempala

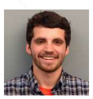
Sampling Problem


Martin Dyer, Alan Frieze, Ravi Kannan

Input: a convex set K with a membership oracle

Output: sample a point from the uniform distribution on K.





Year/Authors	New ingredients	Steps
1989/Dyer-Frieze-Kannan [6]	Everything	n^{23}
1990/Lovász-Simonovits [18]	Better isoperimetry	n^{16}
1990/Lovász [17]	Ball walk	n^{10}
1991/Applegate-Kannan [2]	Logconcave sampling	n^{10}
1990/Dyer-Frieze [5]	Better error analysis	n^{8}
1993/Lovász-Simonovits [19]	Localization lemma	n^{7}
1997/Kannan-Lovï¿œsz-Simonovits [11]	Speedy walk, isotropy	n^{5}
2003/Lovász-Vempala [20]	Annealing, hit-and-run	n^{4}
2015/Cousins-Vempala [3] (well-rounded)	Gaussian Cooling	n^{3}

Theorem: For any convex set, we can sample in $n^{3.5}$ (unconditional) / n^{3} (under KLS conj) steps.

(Same runtime for volume.)

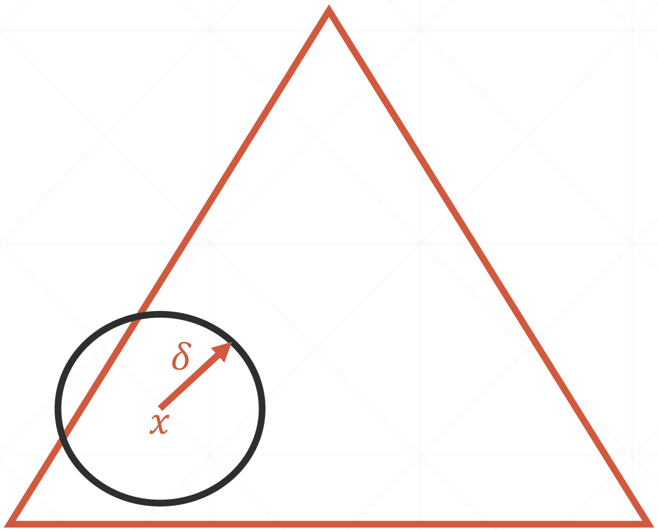
Story Time

Ball Walk

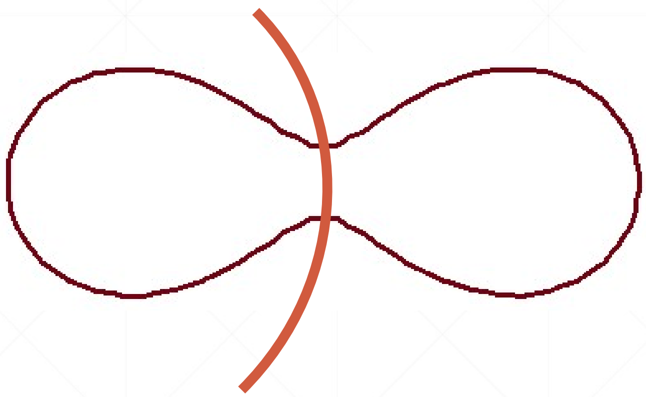
At x, pick random y from $x + \delta B_n$,

if y is in K, go to y.

otherwise, sample again



This walk may get trapped on one side if the set is not convex.



Cheeger constant

For any set K, we define the Cheeger constant ϕ_K by

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

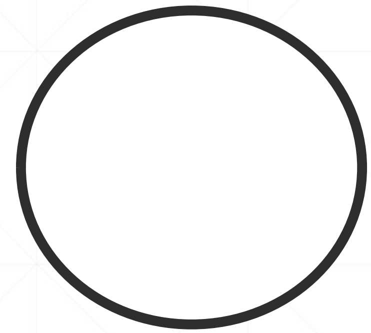
Theorem

Given a random point in K, we can generate another in

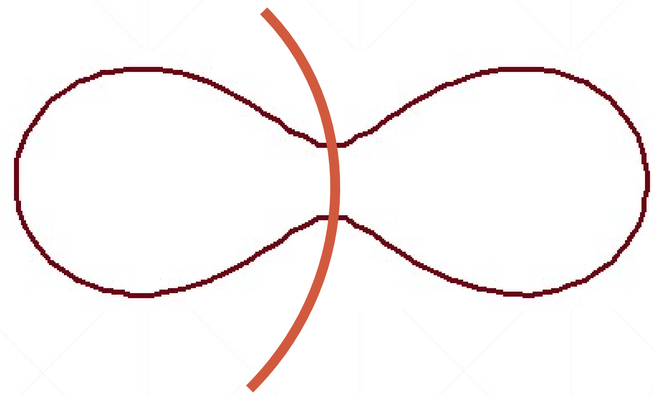
$$O\left(\frac{n}{\delta^2 \phi_K^2} \log(1/\varepsilon)\right)$$

iterations of Ball Walk where δ is step size.

- ϕ_K and δ larger, mix better.
- δ cannot be too large, otherwise, fail probability is ~1.



ϕ large, hard to cut the set



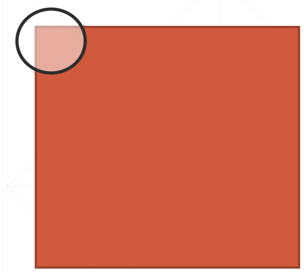
ϕ small, easy to cut the set

Cheeger constant of Convex Set

$\mathrm{Cov}(K) = \mathbb{E}_{x\sim K} x x^T$

Note that ϕ_K is not affine invariant and can be arbitrary small.

$\phi_K = 1/L.$



However, you can renormalize K such that $\mathrm{Cov}(K) = I$.

Definition: K is isotropic, if it is mean 0 and $\mathrm{Cov}(K) = I$.

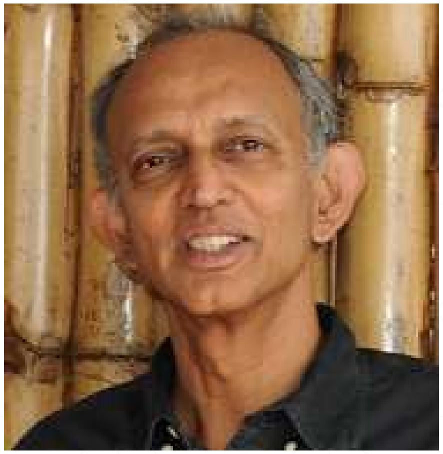
Theorem: If isotropic, $\delta < \frac{0.001}{\sqrt{n}}$, ball walk stays inside the set with constant probability.

Theorem: Given a random point in isotropic K, we can generate another in $O(\frac{n^2}{\phi_K^2} \log(1/\varepsilon))$

KLS Conjecture

Kannan-Lovász-Simonovits Conjecture:

For any isotropic convex K, $\phi_K = \Omega(1)$.



Ravindran Kannan

Lovász László

Miklós Simonovits

Previous Results

[Lovasz-Simonovits 93] $\phi = \Omega(1)n^{-1/2}$.

[Klartag 06] $\sigma = \Omega(1)n^{-1/2}\log^{1/2} n$.

[Fleury-Guedon-Paouris 06] $\sigma = \Omega(1)n^{-1/2}\log^{1/6} n \log^{-2}\log n$.

[Klartag 06] $\sigma = \Omega(1)n^{-0.4}$.

[Fleury 10] $\sigma = \Omega(1)n^{-0.375}$.

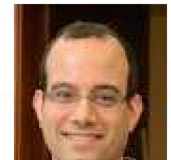
[Guedon-Milman 10] $\sigma = \Omega(1)n^{-0.333}$.

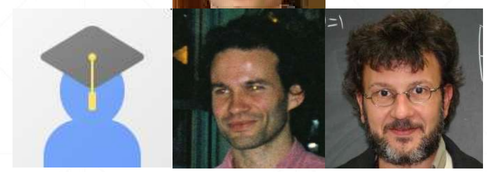
[Eldan 12] $\phi = \widetilde{\Omega}(1)\sigma = \widetilde{\Omega}(1)n^{-0.333}$.

[Lee-Vempala 16] $\phi = \Omega(1)n^{-0.25}$.

What if we cut the body by sphere only?

$$\sigma \stackrel{\text{def}}{=} Var(||X||)^{-1/2} \geq \phi$$



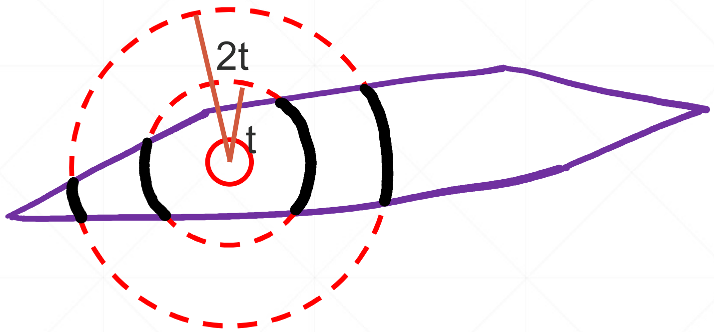


For isotropic convex sets, we can sample in $n^{2.5}$ (unconditional) / n^2 (under KLS) time.

How to make the body isotropic?

Lovász-Vempala Rounding Algorithm

- Start with a ball B inside K
- While B does not cover K
  - Use $O(n)$ samples to estimate the covariance of $K \cap B$.
  - Transform K to make $K \cap B$ isotropic.
  - $B \leftarrow 2B$.



Total Complexity = $\log(n) \cdot n \cdot n^3$.

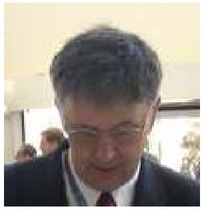
Lemma. $K \cap B$ isotropic $\Rightarrow K \cap 2B$ well-rounded, i.e. $\mathbb{E}||x||^2 = O(n)$ and $\mathrm{Cov}(K) \succcurlyeq \Omega(I)$.

Lemma. We can sample a well-rounded body in time $O(n^3)$ time.

Best known even under KLS conj.

Theorem [Srivastava-Vershynin 13]. M = the empirical covariance of K using n/ϵ^2 samples. Then

$$(1-\epsilon)M \preccurlyeq \mathrm{Cov}(K) \preccurlyeq (1+\epsilon)M$$

Lovász-Vempala at 2006

There is one possible further improvement on the horizon. … If this conjecture is true... could perhaps lead to an $O^*(n^3)$ volume algorithm. But besides the mixing time, a number of further problems concerning achieving isotropic position would have to be solved.

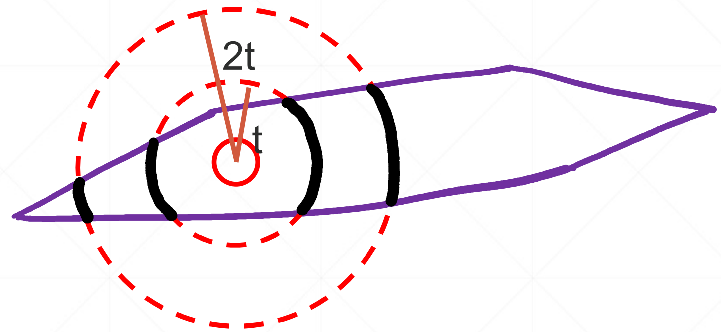
Rounding++

A faster rounding algorithm

How to make the body isotropic?

Lovász-Vempala Rounding Algorithm

- Start with a ball B inside K
- While B does not cover K
  - Use $O(n)$ samples to estimate the covariance of $K \cap B$.
  - Transform K to make $K \cap B$ isotropic.
  - $B \leftarrow 2B$.



Total Complexity = $\log(n) \cdot n \cdot n^3$.

Suffice to make a well-rounded body isotropic.

L ... $x||^2 = O(n)$ and $\mathrm{Cov}(K) \succcurlyeq \Omega(I)$.

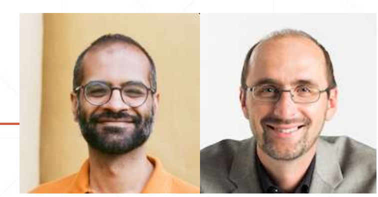
L ... (n^3) time.

Best known even under KLS conj.

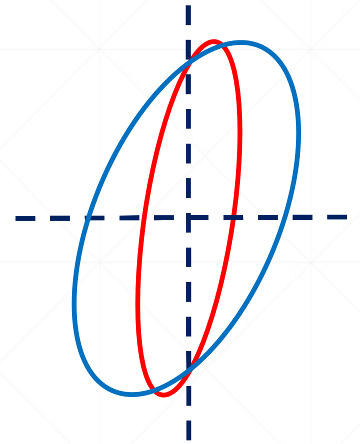
T ... l covariance of K using n/ϵ^2 samples.

T

$$(1-\epsilon)M \preccurlyeq \mathrm{Cov}(K) \preccurlyeq (1+\epsilon)M$$



How to make the well-rounded body isotropic?



Rounding++

- $r \leftarrow 1$
- While $r^2 \leq n$
  - Use $\tilde{O}(r^2)$ samples to estimate the covariance of K.
  - Let V be the subspace of the empirical covariance with eigenvalues $\geq n$. ← $\mathrm{Cov}(K) \preccurlyeq n \cdot I$ ✓
  - Scale up all directions in $V^{\perp}$ by a factor of 2. ← If empirical covariance is accurate, $B(0, r) \subset K$
  - $r \leftarrow 2\left(1 - \frac{1}{\log n}\right) r$. ← We only need $\log(n)$ steps. ✓

Intuition:

- We keep scaling up eigenvalues whenever $\leq n$. So, all eigenvalues converges to n.
- Initially, K far from isotropic. We only need **few expensive** samples.
- At the end, K close to isotropic. We can afford **many cheap** samples.

Why r^2 samples enough to find all eigenvalues $\geq n$?

Lemma [Matrix Chernoff, Ahlswede-Winter]:

A: covariance, $\hat{A}$: empirical covariance of k samples. Then,

$$\hat{A} = (1 \pm \varepsilon)A \pm \tilde{O}\left(\frac{Tr(A)}{\varepsilon k}\right)I.$$

Claim: $\mathrm{Tr}A = O(r^2 n)$.

With $\epsilon = 1/2$ and $k = r^2$, we have $\hat{A} = \left(1 \pm \frac{1}{2}\right) A \pm nI$. Suffices to detect eigenvalues $\geq \Theta(n)$.

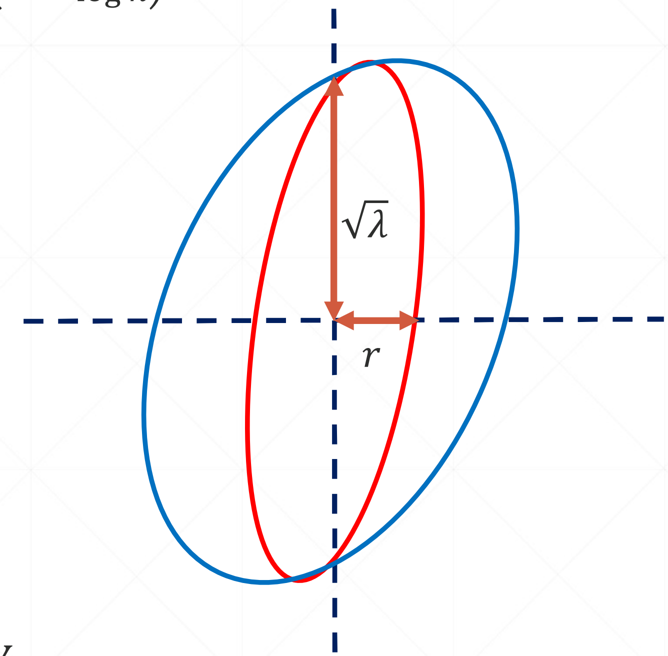
Proof of Claim:

Each step, we scale up some direction by a factor of 2 and $\mathrm{Tr}A$ increased by at most 4.

Since each step r around double, we have $\mathrm{Tr}A = O(r^2 n)$.

$B(0, r) \subset K$

Lemma. While $\lambda \geq 4r^2 \log n$, r increases by a factor of at least $2\left(1 - \frac{1}{\log n}\right)$ in each iteration. (We use $\lambda = n$).



Proof:

Scale up all directions with variance $< \lambda$.

V contains ellipsoid with minimum axis length λ

$V^{\perp}$ contains a ball of radius r that is scaled up by 2.

Then, new body contains a ball of radius nearly $2r$.

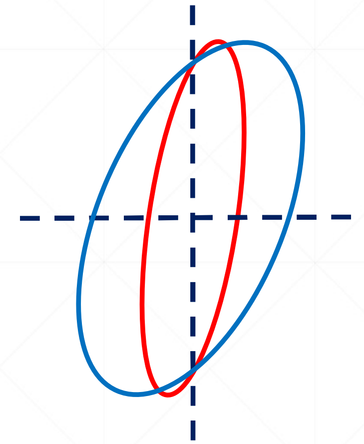
Consider any x on the boundary, we have

$x = \alpha y + (1 - \alpha)$ where $\alpha \in [0,1], y \in \partial B(2r) \cap V^{\perp}, z \in \partial B^n(\lambda) \cap V$

Then,

$$\|x\|^2 = \alpha^2 4r^2 + (1 - \alpha)^2 \lambda \geq \frac{4\lambda r^2}{\lambda + 4r^2} \geq 4 \cdot \frac{\log n}{\log n + 1} \cdot r^2$$

How to make the well-rounded body isotropic?



Rounding++

- $r \leftarrow 1$
- While $r^2 \leq n$
  - Use $\tilde{O}(r^2)$ samples to estimate the covariance of K.
  - Let V be the subspace of the empirical covariance with eigenvalues $\geq n$.
  - Scale up all directions in $V^{\perp}$ by a factor of 2.
  - $r \leftarrow 2\left(1 - \frac{1}{\log n}\right) r$.

$\mathrm{Cov}(K) \preccurlyeq n \cdot I$

If empirical covariance is accurate, $B(0,r) \subset K$

We only need $\log(n)$ steps.

Intuition:

- We keep scaling up eigenv… o n.
- Initially, K far from isotropic…
- At the end, K close to isotropic. We can afford **many cheap** samples.

Under KLS, $\mathrm{Cov}(K) \preccurlyeq n \cdot I$ and $B(0,r) \subset K$
implies
n^3/r^2 time per sample.
So, each phase takes n^3 time.

Without KLS: Isoperimetry for non-isotropic sets

Theorem [Lee-Vempala 16]

$$\phi_K = \Omega(||\mathrm{Cov}K||_F^{-1/2})$$

In particular, $\phi_K = \Omega(n^{-1/4})$ for any isotropic K.

Corollary [This paper] We have complexity $n^{3.5}$.

Lemma [This paper]

Suppose $\phi_K \geq n^{-\beta}$ for isotropic K. For any convex K, we have

$$\phi_K = \widetilde{\Omega}(||\mathrm{Cov}K||_{1/(2\beta)}^{-1/2})$$

(Namely, it suffices to understand isoperimetry for isotropic sets.)

Proof: stochastic localization.

Corollary [This paper] If $\phi_K \geq n^{-\beta}$, we have complexity $n^{3+2\beta}$.

Is $\phi_K \geq n^{-\beta}$ for some $\beta < \frac{1}{4}$?

Extra motivation

- **Theorem (CLT for convex bodies) [Klartag 06]**

For any isotropic log-concave p in $\mathbb{R}^n$,

$$d_{TV}\left(\pi_x p, \mathcal{N}(0,1)\right) \leq o_n(1) \text{ with high prob in } x \sim S^{n-1}$$

Theorem: $W_2\left(p^\top q, \mathcal{N}(0,n)\right) = O(n^{2\beta+\epsilon})$

So, $\beta < \frac{1}{4}$ implies GCLT holds.

- **Conjecture (Generalized CLT for convex bodies)**

Haotian Jiang

For any isotropic log-concave p, q in $\mathbb{R}^n$,

$$d_{TV}\left(\pi_x p, \mathcal{N}(0,n)\right) \leq o_n(1) \text{ with high prob in } x \sim q$$

This version is not symmetric enough. Alternatively:

$$W_2(p^\top q, \mathcal{N}(0,n)) = o_n(\sqrt{n})$$