Coordinate Hit-and-run

Amit Rajaraman

December 22, 2021

Table of Contents

Table of Contents

The problem

Uniformly sampling points from a high-dimensional convex body is a basic problem that relates to problems such as volume computation of convex bodies in high dimensions.

Definition 1

A compact convex body $K \subseteq \mathbb{R}^n$ is said to have a *well-guaranteed membership oracle* if

- it has a *membership oracle*, that is, an oracle that given any $x \in \mathbb{R}^n$ returns whether or not $x \in K$.
- we are given $R > r > 0$ such that $rB_2^n \subseteq K \subseteq RB_2^n$.

Given a well-guaranteed membership oracle for K, the problem is to approximately uniformly sample points for K.

The problem

Question

Input: A convex body $K \subseteq \mathbb{R}^n$ with a well-guaranteed membership oracle.
Output: A probability distribution on $\mathbb{R}^n$ that is at total variation distance at most ϵ from the uniform distribution on K.

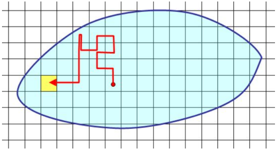
Denote by π_K the uniform distribution on K.

The approach in broad strokes

The primary method to approximately uniformly sample points is through Markov chain Monte Carlo methods.

We synthesize a Markov chain whose stationary distribution is the uniform distribution on the given body and try to determine how fast we get 'close' to this stationary distribution.

The first such random walk that mixes in polynomial time was proposed in [DFK91] by Dyer, Frieze, and Kannan. This random walk was on a grid superimposed on the convex body.



Hit-and-run

The hit-and-run random walk was proved to mix in polynomial time by Lovász in [Lov99].

Definition 2 (Hit-and-run)

Given x_t, we first draw y uniformly at random from $\mathbb{S}^{n-1}$. We then draw x_{t+1} uniformly at random from the set

$$K \cap (x_t + y\mathbb{R}).$$

Since the scheme is ergodic and reversible with respect to π_K, its stationary distribution is π_K.

Coordinate hit-and-run

In the simpler *coordinate hit-and-run* walk (which is the subject of this presentation), we instead draw y uniformly randomly from $\{e_1, \dots, e_n\}$, the standard basis of $\mathbb{R}^n$.
In other words,

Definition 3 (Coordinate hit-and-run)

For the coordinate hit-and-run (CHR) Markov scheme $\mathcal{C}$, the transition probability density from u to v (with respect to the 1-dimensional Lebesgue measure) is

$$\mathcal{C}_{uv} = \begin{cases} \frac{1}{n|K \cap (u + e_i\mathbb{R})|}, & v \in K \cap (u + e_i\mathbb{R}) \text{ for some } i \in [n], \\ 0, & \text{otherwise.} \end{cases}$$

The above is also sometimes referred to as the *Gibbs sampler*.

Warm start

While we have mentioned mixing throughout in a general context, we in fact mean mixing from a 'warm start'.

Definition 4

A distribution P is said to be *M-warm* with respect to another distribution Q if the Radon-Nikodym derivative of P with respect to Q is well-defined and at most M at any point.

In the actual sampling process, we shall bound the mixing time assuming that we are given some distribution that is M-warm with respect to the stationary distribution of the chain.

Conductance

The primary tool we use to determine how fast the Markov chains mix from a warm start is *conductance*.

Definition 5 (Conductance)

Given a Markov scheme P on S, the *ergodic flow* $\Phi_{P,Q}$ with respect to a probability distribution Q on S is

$$\Phi_{P,Q}(A, B) = \int_A P(u, B) Q(\mathrm{d}u)$$

for measurable $A, B \subseteq S$. We also define $\Phi_{P,Q}(A) = \Phi_{P,Q}(A, S \setminus A)$. For $0 < s < 1/2$, the *s-conductance* of P with respect to Q is

$$\Phi_s = \inf_{A : s < Q(A) < 1/2} \frac{\Phi_{P,Q}(A)}{Q(A) - s}.$$

The 0-conductance is referred to as merely *conductance*.

Conductance to mixing

We assert the rapid mixing condition by bounding the s-conductance.

Theorem 6

If a Markov scheme P has stationary distribution Q, and we start it with initial distribution Q_0 with

$$H_s = \sup\{|Q(A) - Q_0(A)| : A \text{ measurable with } Q(A) < s\},$$

then

$$d_{\mathrm{TV}}(Q_t, Q) \leq H_s \left(1 + \frac{1}{s}\left(1 - \frac{\Phi_s^2}{2}\right)^t\right).$$

If Q_0 is M-warm with respect to Q, then $H_s \leq Ms$. Setting $s = \epsilon/(2M)$, we get that

$$d_{\mathrm{TV}}(Q_t, Q) < \epsilon \text{ if } t \geq \log(2M/\epsilon)\Phi_s^{-2}.$$

We give two polynomial bounds on the s-conductance of the CHR scheme.

Two bounds

The first bound is due to [NS21], showing that if $B_\infty^n \subseteq K \subseteq R \cdot B_\infty^n$, then

$$\Phi_{\mathcal{C},s} = \Omega\left(\frac{s^2}{R^2 n^{3.5} (\log n)^3}\right).$$

The second is due to [LV21], showing that if $B_2^n \subseteq K$ and $R_0^2 = \mathbf{E}_{x \sim \pi_K} \|x - b_K\|_2^2$, then

$$\Phi_{\mathcal{C},s} = \Omega\left(\frac{s}{R_0 n^{4.5} \log n}\right).$$

Comparing the two bounds

For any body, R is within a factor of $O(\sqrt{n})$ of R_0. Depending on the body, either of the two bounds can be better. The first bound (featuring R) is better for "cube-like" bodies, and the second bound is better for long thin "tube-like" bodies.

Table of Contents

Overview of the proof

Definition 7 (Gaussian CHR)

Given x_t, we draw i uniformly from $[n]$ and κ from $\mathcal{N}(0, \sigma^2)$. The next point x_{t+1} is defined by

$$x_{t+1} = \begin{cases} x_t + \kappa e_i, & (x_t + \kappa e_i) \in K, \\ x_t, & \text{otherwise.} \end{cases}$$

On bounding the Radon-Nikodym derivative of $\mathcal{C}(x, \cdot)$ with respect to $\mathcal{G}(x, \cdot)$ from below, we arrive at a bound for the conductance of $\mathcal{C}$.

The original strategy used to prove the convergence of usual hit-and-run was to show that 'close-by' points have 'similar' distributions. Here, we do a multistep version of the same, showing that the distributions become similar after running the walk for several steps.

Some setup for the proof

Denote by $\mathcal{G}_v^{(\tau)}$ the probability distribution obtained on starting $\mathcal{G}$ at $v \in \mathbb{R}^n$ and letting it run for τ time steps.

- Let $\mathcal{M}_{n,\tau}$ be the set of all $\mathbb{I} = (i_1, \ldots, i_n) \in \mathbb{N}^n$ with $\sum i_j = \tau$.
- For each such $\mathbb{I}$, define $\mathcal{G}_{v,\mathbb{I}}$ as the Gaussian distribution centered at v with diagonal covariance matrix $\Sigma_{\mathbb{I}}$ defined by $(\Sigma_{\mathbb{I}})_{jj} = \mathbb{I}_j \sigma^2$.
- Finally, define $\lambda_{\mathbb{I}} = \binom{\tau}{\mathbb{I}} n^{-\tau}$, which is the probability of getting the multi-index τ on uniformly drawing an element of $[n]$ τ times.

Some setup for the proof (contd.)

Observe that if we drop the $v \in K$ condition in the definition of $\mathcal{G}$, we just get the random walk $\mathcal{H}$ defined by the mixture

$$\mathcal{H}_v^{(\tau)} = \sum_{\mathbb{I} \in \mathcal{M}_{n,\tau}} \lambda_{\mathbb{I}} \mathcal{G}_{v,\mathbb{I}}.$$

Also define

$$(\mathcal{H}')_v^{(\tau)} = \sum_{\substack{\mathbb{I} \in \mathcal{M}_{n,\tau} \\ \mathbb{I}_j \neq 0 \text{ for all } j}} \lambda_{\mathbb{I}} \mathcal{G}_{v,\mathbb{I}}.$$

Showing the closeness property for $\mathcal{G}$

$$\mathcal{G}_{v,\mathbb{I}} = v + \mathcal{N}(0, \sigma^2 \Sigma_{\mathbb{I}})$$
$$\mathcal{H}_v^{(\tau)} = \sum_{\mathbb{I} \in \mathcal{M}_{n,\tau}} \lambda_{\mathbb{I}} \mathcal{G}_{v,\mathbb{I}}$$

To show that close points have similar distributions, we shall first show that

$$d_{\mathsf{TV}}(\mathcal{G}_v^{(\tau)}, \mathcal{H}_v^{(\tau)}) \text{ is small} \tag{1}$$

and then that for close points u, v,

$$d_{\mathsf{TV}}(\mathcal{H}_u^{(\tau)}, \mathcal{H}_v^{(\tau)}) \text{ is small.} \tag{2}$$

A proof of (1)

$$\mathcal{G}_{v,\mathbb{I}} = v + \mathcal{N}(0, \sigma^2 \Sigma_{\mathbb{I}})$$
$$\mathcal{H}_v^{(\tau)} = \sum_{\mathbb{I} \in \mathcal{M}_{n,\tau}} \lambda_{\mathbb{I}} \mathcal{G}_{v,\mathbb{I}}$$

$$d_{\mathsf{TV}}(\mathcal{G}_v^{(\tau)}, \mathcal{H}_v^{(\tau)}) \text{ is small.}$$

Let $x_0 = v$. For $t \in [\tau]$, pick $i \in [n]$ uniformly randomly, $\kappa \sim \mathcal{N}(0, \sigma^2)$, and let $x_t = x_{t-1} + \kappa e_i$. x_τ just represents a draw from $\mathcal{H}_v^{(\tau)}$.
Consider the event E that every x_i is in K. By a standard coupling argument,

$$d_{\mathsf{TV}}(\mathcal{G}_v^{(\tau)}, \mathcal{H}_v^{(\tau)}) \leq 1 - \Pr[E].$$

Now, set $\tau = 20 n \log n$ and suppose that $\inf_{z \in \partial K} \|v - z\|_\infty > 100 \sigma \log n$. An application of the Chernoff bound (to show that no coordinate is chosen many times) together with a Gaussian tail bound (to show that none of the coordinate changes are too large) shows that with high probability, the points remain in K.

A proof of (2)

$$\mathcal{H}_v^{(\tau)} = \sum_{\mathbb{I} \in \mathcal{M}_{n,\tau}} \lambda_{\mathbb{I}} \mathcal{G}_{v,\mathbb{I}}$$
$$\mathcal{H}'^{(\tau)}_v = \sum_{\substack{\mathbb{I} \in \mathcal{M}_{n,\tau} \\ \forall j, \mathbb{I}_j \neq 0}} \lambda_{\mathbb{I}} \mathcal{G}_{v,\mathbb{I}}$$

$$d_{\mathsf{TV}}(\mathcal{H}_u^{(\tau)}, \mathcal{H}_v^{(\tau)}) \text{ is small.}$$

To show this, first observe that

$$d_{\mathsf{TV}}(\mathcal{H}_u^{(\tau)}, (\mathcal{H}')_u^{(\tau)}) \text{ is small.}$$

Indeed, the distance between the two is at most the sum of all $\lambda_{\mathbb{I}}$ over all $\mathbb{I} \in \mathcal{M}_{n,\tau}$ with a zero in them, and this in turn can be made small for large τ. Therefore, it suffices to show that

$$d_{\mathsf{TV}}((\mathcal{H}')_u^{(\tau)}, (\mathcal{H}')_v^{(\tau)}) \text{ is small.}$$

A proof of (2) (contd.)

For this, we use Pinsker's inequality to get that for any $\mathbb{I} \in \mathcal{M}_{n,\tau}$ with no non-zero element,

$$\begin{aligned} d_{\mathsf{TV}}(\mathcal{G}_{v,\mathbb{I}}, \mathcal{G}_{u,\mathbb{I}}) &\leq \sqrt{\frac{1}{2} D_{\mathsf{KL}}(\mathcal{G}_{v,\mathbb{I}} \| \mathcal{G}_{u,\mathbb{I}})} \\ &= \frac{1}{2}\sqrt{(v-u)^\top \Sigma_{\mathbb{I}}^{-1}(v-u)} \\ &\leq \frac{1}{2\sigma}\|v-u\|_2 \,. \end{aligned}$$

Therefore,

$$d_{\mathsf{TV}}((\mathcal{H}')_u^{(\tau)}, (\mathcal{H}')_v^{(\tau)}) \leq \sum_{\substack{\mathbb{I} \in \mathcal{M}_{n,\tau} \\ \mathbb{I}_j \neq 0 \text{ for any } j}} \lambda_{\mathbb{I}} d_{\mathsf{TV}}(\mathcal{G}_{v,\mathbb{I}}, \mathcal{G}_{u,\mathbb{I}}) \leq \frac{\|v-u\|_2}{2\sigma},$$

which is small.

Concluding the main proof

Now, consider the multistep Gaussian CHR walk $\mathcal{G}'$, performing $\tau = 20n\log n$ steps of $\mathcal{G}$ at each step. By (1) and (2), $\mathcal{G}'$ satisfies the following property. Let

$$K' = \{x \in K : \inf_{z\in\partial K} \|x - z\|_\infty > 100\sigma \log n\}.$$

Then, for any $u, v \in K'$ with $\|u - v\|_2 \leq \sigma$,

$$d_{\mathrm{TV}}(\mathcal{G}'(u,\cdot), \mathcal{G}'(v,\cdot)) \leq 1 - \frac{1}{4}.$$

Given that $B_\infty \subseteq K$, we may further show that $(1 - 100\sigma \log n)K \subseteq K'$. Consequently,

$$\mathrm{vol}(K') \geq (1 - 100\sigma n \log n)\,\mathrm{vol}(K).$$

This is referred to as the (ϵ, δ, ν) property.

Now, how do we go from the (ϵ, δ, ν) property of $\mathcal{G}'$ to a bound on the conductance? Let $S_1 \sqcup S_2$ be a partition of K into disjoint measurable subsets. Let $T_i = S_i \cap K'$ (so we may use the (ϵ, δ, ν) property). Define

$$T_i' = \{x \in T_i : \mathcal{G}'(x, S_{3-i}) < \nu/2\}.$$

Observe that for any $x \in T_1', y \in T_2', \|x - y\|_2 > \delta$. Indeed, otherwise,

$$1 - \nu \geq d_{\mathsf{TV}}(\mathcal{G}'(x, \cdot), \mathcal{G}'(y, \cdot)) \geq 1 - \mathcal{G}'(x, S_2) - \mathcal{G}'(y, S_1) > 1 - \nu.$$

We now use Theorem 2.6 of [LS93].

Theorem 8

Let $\delta > 0$ and $\|\cdot\|_\ell$ be a norm on $\mathbb{R}^n$. Let K be a convex body in $\mathbb{R}^n$ and K_1, K_2 be disjoint measurable subsets of K such that for $u \in K_1$ and $v \in K_2, \|u - v\|_\ell \geq \delta$. Further suppose that $\sup_{x,y \in K} \|x - y\|_\ell = D$. Then,

$$\pi_K(K \setminus (K_1 \cup K_2)) \geq \frac{2\delta}{D - \delta} \min\{\pi_K(K_1), \pi_K(K_2)\}.$$

The previous theorem implies that

$$\pi_K(K' \setminus (T_1' \cup T_2')) \geq \frac{2\delta}{D-\delta} \min\{\pi_K(T_1'), \pi_K(T_2')\}.$$

If $\pi_K(T_1') \leq \frac{1}{2}\pi_K(T_1)$, then the result immediately follows since

$$\Phi_{\mathcal{G}',\pi_K}(S_1, S_2) \geq \Phi_{\mathcal{G}',\pi_K}(T_1 \setminus T_1', S_2) \geq \frac{\nu}{4}(\pi_K(S_1) - \epsilon).$$

Therefore, assume that $\pi_K(T_i') > \frac{1}{2}\pi_K(T_i)$. We now have

$$\begin{aligned}
\Phi_{\mathcal{G}',\pi_K}(S_1, S_2) &\geq \frac{1}{2}(\Phi_{\mathcal{G}',\pi_K}(T_1 \setminus T_1', S_2) + \Phi_{\mathcal{G}',\pi_K}(S_1, T_2 \setminus T_2')) \\
&\geq \frac{\nu}{4}(\pi_K(T_1 \setminus T_1') + \pi_K(T_2 \setminus T_2')) \\
&= \frac{\nu}{4}(\pi_K(K' \setminus (T_1' \cup T_2'))) \\
&\geq \frac{\nu\delta}{2(D-\delta)} \min\{\pi_K(T_1'), \pi_K(T_2')\} \\
&\geq \frac{\nu\delta}{4(D-\delta)}(\min\{\pi_K(S_1), \pi_K(S_2)\} - \epsilon).
\end{aligned}$$

This gives a bound on the ϵ-conductance of $\mathcal{G}'$!

Going from $\mathcal{G}'$ to $\mathcal{G}$

Now, we must arrive at a bound on the conductance of $\mathcal{G}$. Fix some measurable $S \subseteq K$. Let $x_0, x_1, \ldots, x_\tau$ be a walk according to the scheme $\mathcal{G}$, where x_0 is sampled from π_K. Let E denote the event that $x_0 \in S$ and $x_\tau \in K \setminus S$, and E_i the event that $x_i \in S$ and $x_{i+1} \in K \setminus S$. Clearly, $\Pr[E] \leq \sum_i \Pr[E_i]$.
Since each x_i is sampled from π_K (the stationary distribution of $\mathcal{G}$ is π_K), $\Pr[E_i] = \Phi_{\mathcal{G},\pi_K}(S, K \setminus S)$. Further, $\Pr[E] = \Phi_{\mathcal{G}',\pi_K}(S, K \setminus S)$. It follows that

$$\Phi_{\mathcal{G},\pi_K}(S, K \setminus S) \geq \frac{1}{\tau}\Phi_{\mathcal{G}',\pi_K}(S, K \setminus S).$$

Completing the proof

Finally, since the Radon-Nikodym derivative of $\mathcal{C}(x, \cdot)$ with respect to $\mathcal{G}(x, \cdot)$ at $y \neq x$ is bounded below by $\sqrt{2\pi}\sigma/2R$ at any point,

$$\begin{aligned} \Phi_{\mathcal{C},\pi_K}(S, K \setminus S) &\geq \frac{\sqrt{2\pi}\sigma}{2R\tau}\Phi_{\mathcal{G}',\pi_K}(S, K \setminus S) \\ &\geq \frac{\sqrt{2\pi}\sigma}{2R\tau} \cdot \frac{\nu\delta}{4(D-\delta)}(\min\{\pi_K(S), \pi_K(K \setminus S)\} - \epsilon) \end{aligned}$$

Substituting $\tau = 20n \log n$, $\nu = 1/4$, $\delta = \sigma$, $\epsilon = s$, $\sigma = 32s/(n \log n)$ and using the fact that $D \leq R\sqrt{n}$, we get

$$\Phi_{\mathcal{C},\pi_K}(S, K \setminus S) \geq \frac{cs^2}{R^2 n^{3.5} (\log n)^3}(\min\{\pi_K(S), \pi_K(K \setminus S)\} - \epsilon),$$

completing the proof.

Table of Contents

Overview of the proof

In this proof, we reduce the problem of bounded conductance to that of an isoperimetric problem on 'axis-disjoint' subsets.

Definition 9

Sets $S_1, S_2 \subseteq \mathbb{R}^n$ are said to be *axis-disjoint* if for any $i \in [n]$, $(S_1 + e_i\mathbb{R}) \cap S_2 = \emptyset$.

Similar to how in the earlier proof we used the (ϵ, δ, ν) property to arrive at an isoperimetric inequality, here we use the following theorem to get the same.

The main result

Theorem 10

Let K be a convex body satisfying $B_2^n \subseteq K$ and let $R_0^2 = \mathbf{E}_{x \sim \pi_K} \|x - b_K\|_2^2$, where b_K is the centroid of K. Let $S_1, S_2 \subseteq K$ be axis-disjoint. Then, there exists a universal constant c such that for any $\epsilon > 0$,

$$\pi_K(K \setminus (S_1 \cup S_2)) \geq \frac{c\epsilon}{R_0 n^{3.5} \log n}(\min\{\pi_K(S_1), \pi_K(S_2)\} - \epsilon).$$

We denote $K \setminus (S_1 \cup S_2)$ as S_3.
To prove this, we consider a grid of cubes, proving an isoperimetric inequality on each of them separately, then combining these together to give a global isoperimetric inequality.

The final part of the proof

Before proving Theorem 10, let us show how a bound on the conductance follows from it. Suppose $S_1 \sqcup S_2$ is a partition of K into disjoint measurable subsets. Set

$$S_i' = \left\{ x \in S_i : \mathcal{C}(x, S_{3-i}) < \frac{1}{2n} \right\}.$$

We claim that S_1' and S_2' are axis-disjoint. Otherwise, let ℓ be an axis-parallel line intersecting both of them. If $x_i \in \ell \cap S_i'$,

$$\frac{1}{2n} > \mathcal{C}(x_i, S_{3-i}) \geq \frac{1}{n} \frac{|\ell \cap S_i|}{|\ell \cap K|}.$$

However, this gives that $|\ell \cap K| > |\ell \cap S_1| + |\ell \cap S_2|$, which is clearly incorrect since the two are equal.
The remainder of the proof follows near-identically to the final part of the proof of the first bound, except that instead of the (ϵ, δ, ν) property to bound conductance, we use the given theorem regarding axis-disjoint sets.

Isoperimetry on a cube

The following isoperimetric inequality is not too hard to prove and we use it extensively.

Lemma 11

Let C be an axis-aligned cube in $\mathbb{R}^n$. For any axis-disjoint $S_1, S_2 \subseteq C$,

$$\pi_C(C \setminus (S_1 \cup S_2)) \geq \frac{1}{4n \log n} \min\{\pi_C(S_1), \pi_C(S_2)\}.$$

Since we have by definition that

$$\pi_C(C \setminus (S_1 \cup S_2)) \geq 1 - \pi_C(S_1) - \pi_C(S_2),$$

we also have that if $\pi_C(S_1) \leq (2/3)$,

$$\pi_C(C \setminus (S_1 \cup S_2)) \geq \frac{1}{16n \log n} \pi_C(S_1). \tag{3}$$

The lattice

We now move on to the proof of the result. Let S_1, S_2 be axis-disjoint subsets of K. Similar to the first proof, set $K' = (1 - \epsilon/20n)K$, and let $S_i' = S_i \cap K'$. Suppose $\mathrm{vol}(S_1') \leq \mathrm{vol}(S_2')$.
Consider the lattice $\delta\mathbb{Z}^n$, where $\delta = \epsilon/80n\sqrt{n}$. This value of δ is chosen to ensure that any of the cubes that intersect S_i' are contained completely within K.
Define $\mathcal{C}$ to be the set of hypercubes in the lattice that intersect S_1, $\mathcal{C}_1$ as the set of cubes in $\mathcal{C}$ where S_1 takes up at most $2/3$ the volume of the cube, and $\mathcal{C}_2$ as $\mathcal{C} \setminus \mathcal{C}_1$.

The case where $\mathcal{C}_1$ has many cubes

If there is a significant number of cubes in $\mathcal{C}_1$, that is, $\text{vol}(\mathcal{C}_1 \cap S_1) \geq \text{vol}(S_1)/2$, then we can use the previous isoperimetry on each of the cubes to arrive at an overall isoperimetric inequality. That is, we have

$$\begin{aligned}
\pi_K(S_3) &\geq \frac{1}{16n\log n} \sum_{\substack{c \in \mathcal{C}_1 \\ c \cap K' \neq \emptyset}} \pi_K(c \cap S_1) \\
&\geq \frac{1}{16n\log n} \left(\pi_K(\mathcal{C}_1 \cap S_1) - \frac{\epsilon}{2} \right) \\
&\geq \frac{1}{32n\log n} \left(\pi_K(S_1) - \epsilon \right).
\end{aligned}$$

The non-trivial part

Therefore, suppose that $\text{vol}(\mathcal{C}_2 \cap S_1) \geq \text{vol}(S_1)/2$.
The idea is as follows. We first show that the cubes adjacent to $\partial\mathcal{C}_2$ not in $\mathcal{C}_2$ have a significant amount of S_3, then use a known isoperimetric inequality to compare the volume of $\mathcal{C}_2$ to that of $\partial\mathcal{C}_2$.
Combining these two inequalities compares the volume of $\mathcal{C}_2$ (and thus S_1) to that of S_3, which is exactly what we desire.
Set $\mathcal{C}_2'$ to be the set of cubes in $\mathcal{C}_2$ that intersect K'.

The isoperimetric inequality (1)

First, let us show the second part in the overview described on the previous slide. In [KLS95], Kannan, Lovász, and Simonovits show that for any convex body K and measurable $S \subseteq K$,

$$\mathrm{vol}(\partial_K(S)) \geq \frac{\log 2}{R_0} \min\{\mathrm{vol}(S), \mathrm{vol}(K \setminus S)\}, \tag{4}$$

where $\partial_K(S)$ is the boundary of S relative to K.
In particular, since we have

$$\mathrm{vol}(\mathcal{C}_2') \leq \frac{3}{2} \mathrm{vol}(S_1 \cap K') \leq \frac{3}{4} \mathrm{vol}(K'),$$

it follows that

The isoperimetric inequality (2)

$$
\begin{aligned}
\operatorname{vol}(\partial_{K'}(\mathcal{C}_2')) &\geq \frac{\log 2}{R_0} \min\{\operatorname{vol}(\mathcal{C}_2'), \operatorname{vol}(K' \setminus \mathcal{C}_2')\} \\
&\geq \frac{\log 2}{R_0} \min\left\{\operatorname{vol}(\mathcal{C}_2'), \frac{1}{4}\operatorname{vol}(K')\right\} \\
&\geq \frac{\log 2}{R_0} \min\left\{\operatorname{vol}(\mathcal{C}_2 \cap K'), \frac{1}{4}\operatorname{vol}(\mathcal{C}_2 \cap K')\right\} \\
&= \frac{\log 2}{4R_0} \operatorname{vol}(\mathcal{C}_2 \cap K') \\
&\geq \frac{\log 2}{4R_0} \left(\frac{1}{2}\operatorname{vol}(S_1) - \frac{\epsilon}{2}\operatorname{vol}(K)\right) \\
&= \frac{\log 2}{8R_0} \left(\operatorname{vol}(S_1) - \epsilon \operatorname{vol}(K)\right).
\end{aligned}
$$

If we manage to compare $\operatorname{vol}(S_3)$ and $\operatorname{vol}(\partial_{K'}(\mathcal{C}_2'))$, we are done.

Comparing $\mathrm{vol}(S_3)$ and $\mathrm{vol}(\partial_{K'}(\mathcal{C}_2'))$ (1)

Since $\mathcal{C}_2$ is a set of cubes, its boundary consists of facets. Let f be one such facet, c_2 be the cube adjacent to it that is in $\mathcal{C}_2$, and c_1 be the other cube adjacent to it.

Since $\mathrm{vol}(c_2 \cap S_1) \geq (2/3)\,\mathrm{vol}(c_2)$, at least $2/3$ of c_1 is reachable from S_1, and is thus not in S_2.

We then have two cases:

1. If $\mathrm{vol}(c_1 \cap S_1) \leq \mathrm{vol}(c_1)/3$, at least $1/3$ of the $2/3$ fraction earlier (that is not in S_2) is not in S_1 either, and is thus in S_3.
2. If $\mathrm{vol}(c_1)/3 \leq \mathrm{vol}(c_1 \cap S_1) \leq 2\,\mathrm{vol}(c_1)/3$, we can use (3) to conclude that

$$\mathrm{vol}(c_1 \cap S_3) \geq \frac{1}{16 n \log n}\,\mathrm{vol}(c_1 \cap S_1) \geq \frac{1}{48 n \log n}\,\mathrm{vol}(c_1).$$

Comparing $\mathrm{vol}(S_3)$ and $\mathrm{vol}(\partial_{K'}(\mathcal{C}_2))$ (2)

The above says that each cube adjacent to a facet of $\partial_K(\mathcal{C}_2)$ that is not in $\mathcal{C}_2$ must have at least a $1/(48n\log n)$ fraction of S_3. Since each such cube is adjacent to at most $2n$ facets, each facet to contributes to a S_3 volume at least

$$\frac{1}{96n^2\log n}\delta^n.$$

That is,

$$\begin{aligned}
\mathrm{vol}(S_3) &\geq \sum_{f\in\partial_{K'}(\mathcal{C}_2')} \frac{1}{96n^2\log n}\delta^n \\
&= \frac{\delta}{96n^2\log n}\sum_{f\in\partial_{K'}(\mathcal{C}_2')}\delta^{n-1} \\
&= \frac{\delta}{96n^2\log n}\mathrm{vol}(\partial_{K'}(\mathcal{C}_2')).
\end{aligned}$$

Putting the two pieces together

Combining the two equations,

$$\begin{aligned}
\mathrm{vol}(S_3) &\geq \frac{\delta}{96 n^2 \log n} \mathrm{vol}(\partial_{K'}(\mathcal{C}_2')) \\
&\geq \frac{\delta}{96 n^2 \log n} \cdot \frac{\log 2}{8 R_0} \left(\mathrm{vol}(S_1) - \epsilon \, \mathrm{vol}(K)\right) \\
&= \frac{c\delta}{R_0 n^2 \log n} \left(\mathrm{vol}(S_1) - \epsilon \, \mathrm{vol}(K)\right) \\
&= \frac{c'\epsilon}{R_0 n^{3.5} \log n} \left(\mathrm{vol}(S_1) - \epsilon \, \mathrm{vol}(K)\right),
\end{aligned}$$

completing the proof.

Thank you!

References I

Martin Dyer, Alan Frieze, and Ravi Kannan.

A random polynomial-time algorithm for approximating the volume of convex bodies.

J. ACM, 38(1):1–17, January 1991.

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.

László Lovász.

Hit-and-run mixes fast.

Mathematical Programming, 86(3):443–461, Dec 1999.

L. Lovász and M. Simonovits.

Random walks in a convex body and an improved volume algorithm.

Random Structures & Algorithms, 4(4):359–412, 1993.

References II

Aditi Laddha and Santosh S. Vempala.
Convergence of gibbs sampling: Coordinate hit-and-run mixes fast.
In *SoCG*, 2021.

Hariharan Narayanan and Piyush Srivastava.
On the mixing time of coordinate hit-and-run.
Combinatorics, Probability and Computing, page 1–13, 2021.