Random Cayley Graphs

Jonathan Hermon, Sam Thomas

University of British Columbia

Stanford University, 11th May 2020

Random Cayley Graphs

- Let G be a finite group.
- In this talk - usually G is abelian or nilpotent.
- We will actually be considering a sequence $G^{(n)}$ of finite groups with $|G^{(n)}| \to \infty$, and their Cayley graphs w.r.t. random sets of generators.

Random Cayley Graphs

The **Cayley digraph** of a group G with respect to $\mathbf{Z} := [Z_1, ..., Z_k] \subseteq G$ is the graph with vertex set G and edge set

$$\{(g, gz) \mid g \in G,\ z \in \mathbf{Z}\}.$$

The (undirected) **Cayley graph** is given by replacing $\mathbf{Z}$ with $\mathbf{Z} \cup \mathbf{Z}^{-1}$, where $\mathbf{Z}^{-1} := [Z_1^{-1}, ..., Z_k^{-1}]$.

[1] We sample with replacements. Since $k \ll \sqrt{n}$ we can also sample without.
$\ll$ means little o and $\lesssim$ means big O.

Random Cayley Graphs

The **Cayley digraph** of a group G with respect to $\mathbf{Z} := [Z_1, ..., Z_k] \subseteq G$ is the graph with vertex set G and edge set

$$\{(g, gz) \mid g \in G,\ z \in \mathbf{Z}\}.$$

The (undirected) **Cayley graph** is given by replacing $\mathbf{Z}$ with $\mathbf{Z} \cup \mathbf{Z}^{-1}$, where $\mathbf{Z}^{-1} := [Z_1^{-1}, ..., Z_k^{-1}]$.

Throughout, we choose k random elements $Z_1, ..., Z_k$ uniformly at random from G, where $k = k_n$ is a function of $n := |G|$, satisfying $1 \ll \log k \ll \log n$.[1]

We call $Z_1, ..., Z_k$ **generators**, even though they may fail to generate G.

Denote the obtained random Cayley graph by G_k and digraph by $\overrightarrow{G_k}$.

[1]We sample with replacements. Since $k \ll \sqrt{n}$ we can also sample without. $\ll$ means little o and $\lesssim$ means big O.

Random Cayley Graphs

The **Cayley digraph** of a group G with respect to $\mathbf{Z} := [Z_1, ..., Z_k] \subseteq G$ is the graph with vertex set G and edge set

$$\{(g, gz) \mid g \in G,\ z \in \mathbf{Z}\}.$$

The (undirected) **Cayley graph** is given by replacing $\mathbf{Z}$ with $\mathbf{Z} \cup \mathbf{Z}^{-1}$, where $\mathbf{Z}^{-1} := [Z_1^{-1}, ..., Z_k^{-1}]$.

Throughout, we choose k random elements $Z_1, ..., Z_k$ uniformly at random from G, where $k = k_n$ is a function of $n := |G|$, satisfying $1 \ll \log k \ll \log n$.[1]

We call $Z_1, ..., Z_k$ **generators**, even though they may fail to generate G.

Denote the obtained random Cayley graph by G_k and digraph by $\overrightarrow{G_k}$.

All our results hold for both G_k and $\overrightarrow{G_k}$, sometimes with different constants.

Throughout G is not random! Only $\mathbf{Z}$ is random.

[1] We sample with replacements. Since $k \ll \sqrt{n}$ we can also sample without. $\ll$ means little o and $\lesssim$ means big O.

Random Cayley Graphs

Why study random Cayley graphs?

- In the spirit of the legacy of Erdős, this is one instance of the general question:

  "how does a random element of $(\cdot)$ looks/behaves like?"

- Establishing certain universal properties about the random walk (e.g. cutoff) and the geometry that almost all choices of generators satisfy:

The Aldous–Diaconis Conjecture

Figure: Persi Diaconis.

- In 86 Aldous and Diaconis coined the term cutoff and conjectured that: simple random walk on G_k exhibits cutoff (i.e., converges abruptly to equilibrium) for large k, around a time $t = t(|G|, k)$ independent of the algebraic structure of G.

The Aldous–Diaconis Conjecture

Figure: David Aldous

- In 86 Aldous and Diaconis coined the term cutoff and conjectured that: simple random walk on G_k exhibits cutoff (i.e., converges abruptly to equilibrium) for large k, around a time $t = t(|G|, k)$ independent of the algebraic structure of G.

- Confirmed by Dou (92) for k s.t. $\log k \gg \log\log |G|$: cutoff at time $\log_k |G|$.
- Dou & Hildebrand (94) - cutoff for abelian G when $k = \lceil (\log n)^a \rceil$ for $a > 1$ at time $\log_{k/\log n} n$ (where $n := |G|$).

Geometric results - spectral gap

- Alon & Roichman (94) - $\forall \varepsilon \in (0,1)$, $\exists C = C(\varepsilon) > 1$ s.t. for all finite G: G_k is a $1-\varepsilon$ expander w.h.p. when $k > C \log_2 |G|$ - meaning

$$\Phi := \min_{A \subset G : |A| \le |G|/2} \frac{|E(A, A^c)|}{2k|A|} \ge 1 - \varepsilon$$

where $E(A, A^c)$ is the set of edges between A and its complement in G_k.

Geometric results - spectral gap

- Alon & Roichman (94) - $\forall \varepsilon \in (0,1)$, $\exists C = C(\varepsilon) > 1$ s.t. for all finite G: G_k is a $1-\varepsilon$ expander w.h.p. when $k > C \log_2 |G|$ - meaning

$$\Phi := \min_{A \subset G : |A| \le |G|/2} \frac{|E(A, A^c)|}{2k|A|} \ge 1 - \varepsilon$$

  where $E(A, A^c)$ is the set of edges between A and its complement in G_k.

- By Cheeger's ineq. Φ is bounded away form 0 iff the spectral-gap is bounded away form 0

  (where spectral-gap = 2nd smallest eigenvalue of $I - P$,
  where P is the transition matrix of the walk)

Geometric results - spectral gap

- Alon & Roichman (94) - $\exists C > 1$ s.t. for all finite G: G_k is an expander w.h.p. when $k > C \log_2 |G|$.
- H. & Thomas (19) - If G is abelian, the spectral-gap of (SRW on) G_k is at most $C|G|^{-2/k}$ w.p. 1, and

  if $k \geq (2+\delta)d(G)$ it is at least $c|G|^{-2/k}$ w.p. $1 - e^{-c(\delta)k}$, where

  $$d(G) := \text{min size of a set which generates } G.$$

  ($k \geq (1+\delta)d(G)$ suffices if $|G|$ belongs to a density 1 set of $\mathbb{N}$.)

Spectral gap - Open Problems

- Alon & Roichman (94) - $\exists C > 1$ s.t. for all finite G: G_k is an expander w.h.p. when $k > C \log_2 |G|$.
- Open Problem: Does G_k become an expander w.h.p. for smaller values of k if G is not abelian?

Spectral gap - Open Problems

- Alon & Roichman (94) - $\exists C > 1$ s.t. for all finite G: G_k is an expander w.h.p. when $k > C \log_2 |G|$.
- Open Problem: Does G_k become an expander w.h.p. for smaller values of k if G is not abelian?

  E.g., is it enough that $k \geq C \log |G^{\mathrm{ab}}|$?

  Here $G^{\mathrm{ab}} := G/[G, G]$ is the **abelianization** of G, and $[G, G]$ is its commutator (the group generated by $\{ghg^{-1}h^{-1} : g, h \in G\}$).
- Open Problem: For $G = \mathcal{S}_n$ do we get an expander for $k = O(1)$?

  (Helfgott, Seress and Zuk (15): for $k = 2$ w.h.p. G_k is connected and the mixing time is at most $n^3 \log n$.)

Geometric results - diameter

- Shapira and Zuck (18) (improving Marklof and Strömbergsson) - For a sequence of abelian $G^{(n)}$ with fixed $d(G^{(n)}) =: d$ and fixed $k \geq d$:

$$(*) \quad |G^{(n)}|^{-1/k}\mathrm{Diameter}(G_k^{(n)})$$

  converges in distribution as $n \to \infty$ to a non-degenerate distribution (with a semi-explicit description), under some mild conditions.

- El-Baz and Pagano (20) - For $H_{d,p}$, the d-dim **Heisenberg group** of $d \times d$ uni-upper triangular matrices with integer entries mod p, for fixed $k \geq d-1$: same limit as in (*) as $p \to \infty$, with $|G|$ replaced with $|G^{\mathrm{ab}}|$.

Geometric results - diameter

- H. & Thomas (19) - G abelian - If $k \geq (1+\varepsilon)d(G)$ and $k \gg 1$ then w.h.p. the "typical distance" from the identity in G_k is concentrated:

  All but $o(|G|)$ vertices of G_k lie at distance between $R - o(R)$ and $R + o(R)$ from the identity, where $R = R(G) \asymp k|G|^{1/k}$.

- Under mild assumptions R is the minimal radius of a ball in $\mathbb{Z}^k$ of size $\geq |G|$.

- If $k \gtrsim \log|G|$ then $\mathrm{Diameter}(G_k) = R + o(R)$.

Geometric results - diameter

- H. & Thomas (19) - G abelian - If $k \geq (1+\varepsilon)d(G)$ and $k \gg 1$ then w.h.p. the "typical distance" from the identity in G_k is concentrated:

  All but $o(|G|)$ vertices of G_k lie at distance between $R - o(R)$ and $R + o(R)$ from the identity, where $R = R(G) \asymp k|G|^{1/k}$.

- Under mild assumptions R is the minimal radius of a ball in $\mathbb{Z}^k$ of size $\geq |G|$.

- If $k \gtrsim \log |G|$ then $\mathrm{Diameter}(G_k) = R + o(R)$.

- H. & Thomas (19) - similar results hold for the Heisenberg group with $|G|$ above replaced with $|G^{\mathrm{ab}}|$.

TV distance and mixing time - definition

- The total variation (**TV**) distance of two distributions μ and ν on the same finite set G is
$$\|\mu - \nu\|_{\mathsf{TV}} := \frac{1}{2} \sum_{x \in G} |\mu(x) - \nu(x)|.$$

- The TV ε-mixing time of a Markov chain $(X_t)_{t \geq 0}$ is
$$t_{\mathrm{mix}}(\varepsilon) := \inf \left\{ t : \max_x \|\mathbb{P}_x[X_t = \cdot] - \pi\|_{\mathsf{TV}} \leq \varepsilon \right\},$$
where π is the stationary distribution.

Cutoff - definition

- The total variation (**TV**) distance is $\|\mu - \nu\|_{\mathsf{TV}} := \frac{1}{2}\sum_x |\mu(x) - \nu(x)|$.
- The TV ε-mixing time is $t_{\mathrm{mix}}(\varepsilon) := \inf\{t : \max_x \|\mathbb{P}_x[X_t = \cdot] - \pi\|_{\mathsf{TV}} \le \varepsilon\}$, where π is the stationary distribution.
- A sequence of Markov chains exhibits (TV) **cutoff** if the ε-mixing time is asymptotically indep. of ε:

$$\lim_{n\to\infty} t_{\mathrm{mix}}^{(n)}(\varepsilon)/t_{\mathrm{mix}}^{(n)}(1-\varepsilon) = 1, \quad \forall\, 0 < \varepsilon < 1 \tag{1}$$

(the superscript (n) indicates that this is the mixing time of the nth chain).

For a random sequence, we say 'cutoff occurs w.h.p.' if (1) holds in distribution.

Cutoff - definition

- A sequence of MCs exhibits (TV) **cutoff** if the ε-mixing time is asymptotically indep. of ε:

$$\lim_{n\to\infty} t_{\mathrm{mix}}^{(n)}(\varepsilon)/t_{\mathrm{mix}}^{(n)}(1-\varepsilon) = 1, \quad \forall\, 0 < \varepsilon < 1. \tag{2}$$

For a random sequence, say 'cutoff occurs w.h.p.' if (1) holds in distribution.

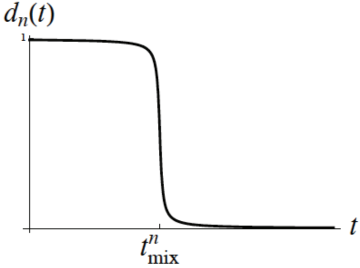

Figure: The worst case TV distance of the nth chain, $d_n(t) := \max_x \|\mathbb{P}_x[X_t = \cdot] - \pi\|_{\mathrm{TV}}$ as a function of t when cutoff occurs.

Universality of cutoff

- Many families of Markov chains are believed to exhibit cutoff, but only few examples are fully understood.
- A recurring theme is that random instances exhibit cutoff.
- Often the cutoff time is an '**entropic time**', meaning a time at which an auxiliary walk, often on the Benjamini-Schramm limit, has the same entropy as the stationary distribution.

Universality of cutoff - random walk on random graphs

Random walk on the following random instances exhibits cutoff at an entropic time:

- Lubetzky and Sly (11) - Random d regular graphs.
- Berestycki, Lubetzky, Peres, Sly (16) - The giant component of an Erdős-Renyi graph and the configuration model with min. degree ≥ 3.
- Bordenave, Caputo and Salez (18) - random digraphs with given degree seq.
- Bordenave and Lacoin (19) - random n-lift of a graph.[2]
- H., Sly, Sousi (20+) - 'quasi-random graphs' - obtained by adding to an arbitrary bounded degree graph (with connected components of size ≥ 3) the edges of a random perfect matching of the vertices.

[2] A random n lift of (V, E) is generated by taking n copies $v_1, \dots, v_n$ of each $v \in V$ and for each $uv \in E$ connect u_i with $v_{\tau_{uv}(i)}$, where τ_{uv} is a random permutation of $[n]$.

Cutoff - results

We prove cutoff at an entropic time for:

- Abelian G when $k - d(G) \gg 1$.[3]

  The entropic time is usually the time that the random walk on $\mathbb{Z}^k$ is $\log |G|$.

[3] Other than when $\sqrt{\log |G| / \log \log \log |G|} \lesssim k \lesssim \sqrt{\log n}$, where we need $k - d(G) \gg \log \log k$.

[4] Throughout $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.

Cutoff - results

We prove cutoff at an entropic time for:

- Abelian G when $k - d(G) \gg 1$.[3]

  The entropic time is usually the time that the random walk on $\mathbb{Z}^k$ is $\log |G|$.

  This time is only a function of k and $|G|$ in accordance to the Aldous-Diaconis conjecture!

  When k is of order close to $\log |G|$ the entropic time will be defined w.r.t. random walk on[4] $\mathbb{Z}_m^k$ for various values of m.

[3] Other than when $\sqrt{\log |G| / \log\log\log |G|} \lesssim k \lesssim \sqrt{\log n}$, where we need $k - d(G) \gg \log\log k$.

[4] Throughout $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.

Cutoff - results

We prove cutoff at an entropic time for:

- Abelian G when $k - d(G) \gg 1$.[3]

  The entropic time is usually the time that the random walk on $\mathbb{Z}^k$ is $\log |G|$.

  This time is only a function of k and $|G|$ in accordance to the Aldous-Diaconis conjecture!

  When k is of order close to $\log |G|$ the entropic time will be defined w.r.t. random walk on[4] $\mathbb{Z}_m^k$ for various values of m.

- The Heisenberg group $H_{p,d}$ of $d \times d$ uni-upper triangular matrices with entries mod a prime p, as $p \to \infty$ (if d is constant or diverges slowly).

  Here the entropic time is the time that the random walk on $\mathbb{Z}^k$ is $\log |G^{\mathrm{ab}}|$.

[3] Other than when $\sqrt{\log |G| / \log\log\log |G|} \lesssim k \lesssim \sqrt{\log n}$, where we need $k - d(G) \gg \log\log k$.

[4] Throughout $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.

Cutoff - results

We prove cutoff at an entropic time for:

- Abelian G when $k - d(G) \gg 1$.[3]

  The entropic time is usually the time that the random walk on $\mathbb{Z}^k$ is $\log |G|$.

  This time is only a function of k and $|G|$ in accordance to the Aldous-Diaconis conjecture!

  When k is of order close to $\log |G|$ the entropic time will be defined w.r.t. random walk on[4] $\mathbb{Z}_m^k$ for various values of m.

- The Heisenberg group $H_{p,d}$ of $d \times d$ uni-upper triangular matrices with entries mod a prime p, as $p \to \infty$ (if d is constant or diverges slowly).

  Here the entropic time is the time that the random walk on $\mathbb{Z}^k$ is $\log |G^{\mathrm{ab}}|$. This time does depend on d, not only on k and $|H_{p,d}|$.

---

[3] Other than when $\sqrt{\log |G| / \log \log \log |G|} \lesssim k \lesssim \sqrt{\log n}$, where we need $k - d(G) \gg \log \log k$.

[4] Throughout $\mathbb{Z}_m := \mathbb{Z}/m\mathbb{Z}$.

Generating the walk via an auxiliary random walk

- One way of generating the walk $S(t)$ on G_k is to take a random word $W(t)$ in the free group F_k, or in the abelian setup the free abelian group $\mathbb{Z}^k$, and then substitute the generators of the free group in W by Z.
- Think of W as the sequence of the indices of the generators picked at each step and their sign (indicating if we picked Z_i or Z_i^{-1}).
- In the abelian setup, $W_i(t)$ is counting how many times Z_i was picked by time t minus how many times $-Z_i$ was picked.

  Here $W(t)$ is a simple random walk on $\mathbb{Z}^k$.

A neat expression for the L_2 distance in our setup

- For a probability measure μ on G the L_2 distance from the uniform distribution π is defined as

$$\|\mu - \pi\|_{2,\pi}^2 := |G| \sum_{g \in G} \left(\mu(g) - \frac{1}{|G|} \right)^2$$

$$\geq^{\text{(by Cauchy-Schwarz)}} \left(\sum_{g \in G} |\mu(g) - \frac{1}{|G|}| \right)^2 = 4\|\mu - \pi\|_{\mathrm{TV}}^2.$$

Observe that $\|\mu - \pi\|_{2,\pi}^2 + 1 := |G| \sum_{g \in G} \mu(g)^2$

A neat expression for the L_2 distance in our setup

- Recall $\|\mu - \pi\|_{2,\pi}^2 + 1 = |G| \sum_{g \in G} \mu(g)^2$.
- If $S(t)$ is our random walk on G_k at time t, and $S'(t)$ is an independent copy (given $Z := [Z_1, \dots, Z_k]$; i.e., given the graph) then

$$\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2 + 1 = |G| \sum_{g \in G} \mathbb{P}[S(t) = g \mid Z]^2$$

$$|G| \sum_{g \in G} \mathbb{P}[S(t) = g = S'(t) \mid Z] = |G| \mathbb{P}[S(t) = S'(t) \mid Z].$$

- We generate S and S' by picking independent walks W, W' on the free group, and then substituting the generators of the free group by Z.

A neat expression for the L_2 distance in our setup

- Recall $\|\mu - \pi\|_{2,\pi}^2 + 1 = |G| \sum_{g \in G} \mu(g)^2$.
- If $S(t)$ is our random walk on G_k at time t, and $S'(t)$ is an independent copy (given $Z := [Z_1, \dots, Z_k]$; i.e., given the graph) then

$$\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2 + 1 = |G| \sum_{g \in G} \mathbb{P}[S(t) = g \mid Z]^2$$

$$|G| \sum_{g \in G} \mathbb{P}[S(t) = g = S'(t) \mid Z] = |G| \mathbb{P}[S(t) = S'(t) \mid Z].$$

- We generate S and S' by picking independent walks W, W' on the free group, and then substituting the generators of the free group by Z.
- Now take expectation.

Transforming a quenched problem to an annealed expectation:

- If $S(t)$ is our walk on G_k at time t, and $S'(t)$ is an independent copy given $Z := [Z_1, \dots, Z_k]$, then

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2] = |G|\mathbb{P}[S(t) = S'(t)] - 1.$$

  Crucially, the r.h.s. is an **annealed** probability (averaging over Z)!

Transforming a quenched problem to an annealed expectation:

- If $S(t)$ is our walk on G_k at time t, and $S'(t)$ is an independent copy given $Z := [Z_1, \ldots, Z_k]$, then

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2] = |G|\mathbb{P}[S(t) = S'(t)] - 1.$$

  Crucially, the r.h.s. is an **annealed** probability (averaging over Z)!

- Hence if r.h.s. is $o(1)$, by Markov's ineq. we can derive (the quenched statement) that w.h.p. Z is such that at time t this distance is $o(1)$.

Transforming a quenched problem to an annealed expectation:

- If $S(t)$ is our walk on G_k at time t, and $S'(t)$ is an independent copy given $Z := [Z_1, \dots, Z_k]$, then

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2] = |G|\mathbb{P}[S(t) = S'(t)] - 1.$$

  Crucially, the r.h.s. is an **annealed** probability (averaging over Z)!

- Hence if r.h.s. is $o(1)$, by Markov's ineq. we can derive (the quenched statement) that w.h.p. Z is such that at time t this distance is $o(1)$.
- This is often enough for determining the order of the mixing time, but too coarse for proving cutoff.

From a quenched problem to an annealed one:

- If $S(t)$ is our walk on G_k at time t, and $S'(t)$ is an independent copy given $Z := [Z_1, \dots, Z_k]$, then

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2] = |G|\mathbb{P}[S(t) = S'(t)] - 1.$$

- Hence if r.h.s. is $o(1)$, by Markov's ineq. we can derive (the quenched statement) that w.h.p. Z is such that at time t this distance is $o(1)$.
- To prove cutoff we do a modified L_2 calculation, by conditioning the words W, W' in the free group, used to generate S and S', to both satisfy some condition that holds w.h.p.

From a quenched problem to an annealed one:

- If $S(t)$ is our walk on G_k at time t, and $S'(t)$ is an independent copy given $Z := [Z_1, \dots, Z_k]$, then

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{2,\pi}^2] = |G|\mathbb{P}[S(t) = S'(t)] - 1.$$

- Hence if r.h.s. is $o(1)$, by Markov's ineq. we can derive (the quenched statement) that w.h.p. Z is such that at time t this distance is $o(1)$.
- To prove cutoff we do a modified L_2 calculation, by conditioning the words W, W' in the free group, used to generate S and S', to both satisfy some condition that holds w.h.p.

  E.g., if some generator is picked only once in W and 0 times in W' or vice versa, then $S^{-1}S' \sim \mathrm{Unif}(G) \implies \mathbb{P}[S(t) = S'(t) \mid \text{this event}] = 1/|G|$.

  (On this event can write $S^{-1}S' = AZ_i^{\pm 1}B$ for A, B indep. of Z_i.)

A modified L_2 bound

We use the modified L_2 bound:

$$\frac{1}{2}\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{\mathrm{TV}}]^2$$

$$\leq |G|\mathbb{P}[S(t) = S'(t) \mid W(t), W'(t) \in \mathrm{typ}] - 1 + \mathbb{P}[W(t) \notin \mathrm{typ}],$$

where typ is a certain 'typical' event (i.e. $\mathbb{P}[W(t) \notin \mathrm{typ}] = o(1)$)

(i.e., in terms of the sequence of indices picked by time t (with multiplicities): $i_1, \ldots, i_{r(t)}$, and their signs).

A modified L_2 bound

We use the modified L_2 bound:

$$\frac{1}{2}\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{\mathrm{TV}}]^2$$

$$\leq |G|\mathbb{P}[S(t) = S'(t) \mid W(t), W'(t) \in \mathrm{typ}] - 1 + \mathbb{P}[W(t) \notin \mathrm{typ}],$$

where typ is a certain 'typical' event (i.e. $\mathbb{P}[W(t) \notin \mathrm{typ}] = o(1)$)

(i.e., in terms of the sequence of indices picked by time t (with multiplicities): $i_1, \ldots, i_{r(t)}$, and their signs).

Proof:

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{\mathrm{TV}}]$$

$$\leq \mathbb{P}[W(t) \notin \mathrm{typ}] + \mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z, W(t) \in \mathrm{typ}] - \pi\|_{\mathrm{TV}}]$$

A modified L_2 bound

We use the modified L_2 bound:

$$\frac{1}{2}\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{\mathrm{TV}}]^2$$

$$\leq |G|\mathbb{P}[S(t) = S'(t) \mid W(t), W'(t) \in \mathrm{typ}] - 1 + \mathbb{P}[W(t) \notin \mathrm{typ}],$$

where typ is a certain 'typical' event (i.e. $\mathbb{P}[W(t) \notin \mathrm{typ}] = o(1)$)

(i.e., in terms of the sequence of indices picked by time t (with multiplicities): $i_1, \ldots, i_{r(t)}$, and their signs).

Proof:

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z] - \pi\|_{\mathrm{TV}}]$$

$$\leq \mathbb{P}[W(t) \notin \mathrm{typ}] + \mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid Z, W(t) \in \mathrm{typ}] - \pi\|_{\mathrm{TV}}]$$

$$\mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid W(t) \in \mathrm{typ}, Z] - \pi\|_{\mathrm{TV}}]^2$$

$$(\text{Cauchy-Schwarz}) \leq \mathbb{E}[\|\mathbb{P}[S(t) = \cdot \mid W(t) \in \mathrm{typ}, Z] - \pi\|_{2,\pi}^2]$$

$$= |G|\mathbb{P}[S(t) = S'(t) \mid W(t), W'(t) \in \mathrm{typ}] - 1. \qquad \square$$

Warm up - a proof of Dou's result: For $\log k \gg \log\log n$ cutoff at time $\log_{2k} n$ (where $n := |G|$)

Proof: $t := \log_{2k} n$ is always a lower bound on the mixing time since in $r = t(1-\varepsilon)$ steps the walk can only see $(2k)^r = o(n)$ vertices.

Warm up - a proof of Dou's result: For $\log k \gg \log\log n$ cutoff at time $\log_{2k} n$ (where $n := |G|$)

Proof: $t := \log_{2k} n$ is always a lower bound on the mixing time since in $r = t(1-\varepsilon)$ steps the walk can only see $(2k)^r = o(n)$ vertices.

Upper bound: Let $k \geq (\log n)^2$. For all choice of typ

$$n\mathbb{P}[S = S' \mid W, W' \in \mathrm{typ}] - 1 \leq n\rho,$$

$\rho := \mathbb{P}[$no generator picked once in W and 0 in W' or vice-versa $\mid W, W' \in \mathrm{typ}]$.

We want a smart choice of typ so that $\rho := o(1/n)$.

Warm up - a proof of Dou's result: For $\log k \gg \log\log n$ cutoff at time $\log_{2k} n$ (where $n := |G|$)

Proof: $t := \log_{2k} n$ is always a lower bound on the mixing time since in $r = t(1-\varepsilon)$ steps the walk can only see $(2k)^r = o(n)$ vertices.

Upper bound: Let $k \geq (\log n)^2$. For all choice of typ

$$n\mathbb{P}[S = S' \mid W, W' \in \mathrm{typ}] - 1 \leq n\rho,$$

$\rho := \mathbb{P}[\text{no generator picked once in } W \text{ and } 0 \text{ in } W' \text{ or vice-versa} \mid W, W' \in \mathrm{typ}]$.

We want a smart choice of typ so that $\rho := o(1/n)$.

We take typ to be the event that each generator is picked at most once, and between $(1+\varepsilon/2)t$ and $s := (1+\varepsilon)t$ generators are used once by time s.

An easy calculation (involving binomial co-efficients and Stirling's approximation) shows that this choice works when $t = \log_{k/\log n} n$

Warm up - a proof of Dou's result: For $\log k \gg \log\log n$ cutoff at time $\log_{2k} n$ (where $n := |G|$)

Proof: $t := \log_{2k} n$ is always a lower bound on the mixing time since in $r = t(1-\varepsilon)$ steps the walk can only see $(2k)^r = o(n)$ vertices.

Upper bound: Let $k \geq (\log n)^2$. For all choice of typ

$$n\mathbb{P}[S = S' \mid W, W' \in \mathrm{typ}] - 1 \leq n\rho,$$

$\rho := \mathbb{P}[$no generator picked once in W and 0 in W' or vice-versa $\mid W, W' \in \mathrm{typ}]$.

We want a smart choice of typ so that $\rho := o(1/n)$.

We take typ to be the event that each generator is picked at most once, and between $(1+\varepsilon/2)t$ and $s := (1+\varepsilon)t$ generators are used once by time s.

An easy calculation (involving binomial co-efficients and Stirling's approximation) shows that this choice works when $t = \log_{k/\log n} n \approx \log_{2k} n$ in Dou's setup.

Warm up - a proof of Dou's result: For $\log k \gg \log\log n$ cutoff at time $\log_{2k} n$ (where $n := |G|$)

Proof: $t := \log_{2k} n$ is always a lower bound on the mixing time since in $r = t(1-\varepsilon)$ steps the walk can only see $(2k)^r = o(n)$ vertices.

Upper bound: Let $k \geq (\log n)^2$. For all choice of typ

$$n\mathbb{P}[S = S' \mid W, W' \in \mathrm{typ}] - 1 \leq n\rho,$$

$\rho := \mathbb{P}[$no generator picked once in W and 0 in W' or vice-versa $\mid W, W' \in \mathrm{typ}]$.

We want a smart choice of typ so that $\rho := o(1/n)$.

We take typ to be the event that each generator is picked at most once, and between $(1+\varepsilon/2)t$ and $s := (1+\varepsilon)t$ generators are used once by time s.

An easy calculation (involving binomial co-efficients and Stirling's approximation) shows that this choice works when $t = \log_{k/\log n} n \approx \log_{2k} n$ in Dou's setup.

A small modification to typ extends the upper bound $\log_{k/\log n} n$ to all $k \gg \log n$. For abelian groups, we can prove a matching lower bound of $\log_{k/\log n} n$ via entropic considerations, thus establishing cutoff for $k \gg \log n$.

Warm up - a proof of Alon-Roichman's result

Proof: When $k \geq C \log_2 n$ let $t = C' \log n$ and pick typ to be the event that roughly the right number of generators are picked once and zero times.

[5] The choice of C' depends on C. The larger C is, the smaller C' is.

Warm up - a proof of Alon-Roichman's result

Proof: When $k \geq C \log_2 n$ let $t = C' \log n$ and pick typ to be the event that roughly the right number of generators are picked once and zero times.

Then $\mathbb{P}[W \notin \mathrm{typ}] \leq n^{-2c}$ and for some choice of[5] C' also $\rho \leq n^{-1-2c}$.

$\implies$ expected (w.r.t. Z) TV distance at time t is at most $2n^{-2c}$,

$$\implies \quad t_{\mathrm{mix}}(1/n^c) \leq C' \log n$$

(with failure prob. $\leq 2n^{-c}$).

[5] The choice of C' depends on C. The larger C is, the smaller C' is.

Warm up - a proof of Alon-Roichman's result

Proof: When $k \geq C \log_2 n$ let $t = C' \log n$ and pick typ to be the event that roughly the right number of generators are picked once and zero times.

Then $\mathbb{P}[W \notin \mathrm{typ}] \leq n^{-2c}$ and for some choice of[5] C' also $\rho \leq n^{-1-2c}$.

$$\implies \text{expected (w.r.t. } Z\text{) TV distance at time } t \text{ is at most } 2n^{-2c},$$

$$\implies \quad t_{\mathsf{mix}}(1/n^c) \leq C' \log n$$

(with failure prob. $\leq 2n^{-c}$).

Conclude using

$$\frac{1}{\mathsf{gap}} \leq \frac{t_{\mathsf{mix}}(\delta/2)}{\log(1/\delta)}$$

[5] The choice of C' depends on C. The larger C is, the smaller C' is.

Warm up - a proof of Alon-Roichman's result

Proof: When $k \geq C \log_2 n$ let $t = C' \log n$ and pick typ to be the event that roughly the right number of generators are picked once and zero times.

Then $\mathbb{P}[W \notin \mathrm{typ}] \leq n^{-2c}$ and for some choice of[5] C' also $\rho \leq n^{-1-2c}$.

$$\implies \text{expected (w.r.t. } Z\text{) TV distance at time } t \text{ is at most } 2n^{-2c},$$

$$\implies \quad t_{\mathsf{mix}}(1/n^c) \leq C' \log n$$

(with failure prob. $\leq 2n^{-c}$).

Conclude using

$$\frac{1}{\mathsf{gap}} \leq \frac{t_{\mathsf{mix}}(\delta/2)}{\log(1/\delta)} \lesssim 1 \text{ for } \delta = 2/n^c.$$

[5] The choice of C' depends on C. The larger C is, the smaller C' is.

Previous Work

- Wilson (97) - Proved cutoff for $\mathbb{Z}_2^d$ conditioned on generating the group, and conjectured that if $|G| \leq 2^d$ then for all k (w.h.p.) G_k has a smaller mixing time (up to smaller order terms) than H_k for $H = \mathbb{Z}_2^d$.
- Hough (17) cutoff for the cyclic group $\mathbb{Z}_p$ with p prime and $1 \ll k \leq \log p / \log \log p$.

Wilson's conjecture

Theorem

Wilson's conjecture is true - $\mathbb{Z}_2^d$ is the "slowest mixing group".

Idea: We work with the lazy random walk, which at each step stays put w.p. $1/2$. The extra randomness coming from the laziness will allow us to condition on Z and W and still keep the walk $S = S(t)$ 'random enough':

Let $g_1, \ldots, g_\ell \in G$ and $\xi_1, \ldots, \xi_\ell$ i.i.d. each equal 0 w.p. $1/2$ and o.w. ± 1 with equal probability.

Wilson's conjecture

Theorem

Wilson's conjecture is true - $\mathbb{Z}_2^d$ is the "slowest mixing group".

Idea: We work with the lazy random walk, which at each step stays put w.p. $1/2$. The extra randomness coming from the laziness will allow us to condition on Z and W and still keep the walk $S = S(t)$ 'random enough':

Let $g_1, \ldots, g_\ell \in G$ and $\xi_1, \ldots, \xi_\ell$ i.i.d. each equal 0 w.p. $1/2$ and o.w. ± 1 with equal probability. 'Coupon collector for groups':

Step 1 (inspired by Erdös & Rényi): The law $g_1^{\xi_1} \cdots g_\ell^{\xi_\ell}$, given $|\sum_i |\xi_i| - \ell/2| = O(\sqrt{\ell})$, is close in TV distance to uniform, for most choice $g_1, \ldots, g_\ell \in G$, provided $\ell \geq \log_2 n - o(\sqrt{n})$.

Wilson's conjecture

Theorem

Wilson's conjecture is true - $\mathbb{Z}_2^d$ is the "slowest mixing group".

Idea: We work with the lazy random walk, which at each step stays put w.p. $1/2$. The extra randomness coming from the laziness will allow us to condition on Z and W and still keep the walk $S = S(t)$ 'random enough':

Let $g_1, \ldots, g_\ell \in G$ and $\xi_1, \ldots, \xi_\ell$ i.i.d. each equal 0 w.p. $1/2$ and o.w. ± 1 with equal probability. 'Coupon collector for groups':

Step 1 (inspired by Erdös & Rényi): The law $g_1^{\xi_1} \cdots g_\ell^{\xi_\ell}$, given $|\sum_i |\xi_i| - \ell/2| = O(\sqrt{\ell})$, is close in TV distance to uniform, for most choice $g_1, \ldots, g_\ell \in G$, provided $\ell \geq \log_2 n - o(\sqrt{n})$.

Step 2 (inspired by Pak): Prove a 'censoring inequality', saying that (in the lazy setup) eliminating from S repetitions of the same generator cannot increases the expected TV distance.

Wilson's conjecture

Theorem

Wilson's conjecture is true - $\mathbb{Z}_2^d$ is the "slowest mixing group".

Idea: We work with the lazy random walk, which at each step stays put w.p. $1/2$. The extra randomness coming from the laziness will allow us to condition on Z and W and still keep the walk $S = S(t)$ 'random enough':

Let $g_1, \ldots, g_\ell \in G$ and $\xi_1, \ldots, \xi_\ell$ i.i.d. each equal 0 w.p. $1/2$ and o.w. ± 1 with equal probability. 'Coupon collector for groups':

Step 1 (inspired by Erdös & Rényi): The law $g_1^{\xi_1} \cdots g_\ell^{\xi_\ell}$, given $|\sum_i |\xi_i| - \ell/2| = O(\sqrt{\ell})$, is close in TV distance to uniform, for most choice $g_1, \ldots, g_\ell \in G$, provided $\ell \geq \log_2 n - o(\sqrt{n})$.

Step 2 (inspired by Pak): Prove a 'censoring inequality', saying that (in the lazy setup) eliminating from S repetitions of the same generator cannot increases the expected TV distance.

Entropic times

- Let $t_0 = t_0(n,k)$ be the time at which the entropy of a rate-1 SRW on $\mathbb{Z}^k$ equals $\log n$.

Entropic times

- Let $t_0 = t_0(n,k)$ be the time at which the entropy of a rate-1 SRW on $\mathbb{Z}^k$ equals $\log n$.
- $t_0 \asymp kn^{2/k}$ when $k \ll \log n$,
- $t_0 \asymp \log n$ when $k \asymp \log n$.
- Let $t_m = t_0(G,k,m)$ be the time at which the entropy of SRW on $\mathbb{Z}_m^k$ becomes $\log|G/mG|$,

$$\text{where for } m \in \mathbb{N},\ mG := \{mg : g \in G\},$$

$$\Big(\text{If } G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d},\ \text{then } mG \cong \mathbb{Z}_{m_1/\gcd(m_1,m)} \oplus \cdots \oplus \mathbb{Z}_{m_d/\gcd(m_d,m)}\Big)$$

- Let

$$T := t_0 \vee \max_{m:\, m|n} t_m.$$

(Remark: $T \asymp t_0$ when $k \geq (1+\delta)d(G)$.)

Our results in the abelian setup

Recall that $1 \ll \log k \ll \log n$ and $n := |G|$. Let $d = d(G)$ be the size of the smallest set of generators.

Theorem (abelian cutoff)

For G abelian, SRW on G_k, exhibits cutoff at the entropic time T provided $k - d(G) \gg 1$

Our results in the abelian setup

Recall that $1 \ll \log k \ll \log n$ and $n := |G|$. Let $d = d(G)$ be the size of the smallest set of generators.

Theorem (abelian cutoff)

For G abelian, SRW on G_k, exhibits cutoff at the entropic time T provided $k - d(G) \gg 1$, other than when $\sqrt{\log |G| / \log\log\log |G|} \lesssim k \lesssim \sqrt{\log n}$, where (only if $|G|$ belongs to a density 0 set of $\mathbb{N}$) we need the slightly stronger assumption $k - d(G) \gg \log\log k$.

Our results in the abelian setup

Recall that $1 \ll \log k \ll \log n$ and $n := |G|$. Let $d = d(G)$ be the size of the smallest set of generators.

Theorem (abelian cutoff)

For G abelian, SRW on G_k, exhibits cutoff at the entropic time T provided $k - d(G) \gg 1$, other than when $\sqrt{\log |G| / \log \log \log |G|} \lesssim k \lesssim \sqrt{\log n}$, where (only if $|G|$ belongs to a density 0 set of $\mathbb{N}$) we need the slightly stronger assumption $k - d(G) \gg \log \log k$.

Under mild conditions $T = t_0$, which depends only on $|G|$ and k, which is consistent with the Aldous-Diaconis conjecture.

The cutoff shape is Gaussian and is governed by the fluctuations of the r.v. whose mean is the entropy of W_t: i.e. $-\log \mu(W_t)$, where μ is the law of W_t.

Heisenberg Matrix Groups

Let $G = H_{p,d}$ be the Heisenberg group of $d \times d$ uni-upper triangular matrices with integer entries mod p.

For $A, B \in G$ we have $(AB)_{i,i+1} = A_{i,i+1} + B_{i,i+1}$ and the Abelianization $G/[G, G]$ is $\cong \mathbb{Z}_p^{d-1}$.

Heisenberg Matrix Groups

Let $G = H_{p,d}$ be the Heisenberg group of $d \times d$ uni-upper triangular matrices with integer entries mod p.

For $A, B \in G$ we have $(AB)_{i,i+1} = A_{i,i+1} + B_{i,i+1}$ and the Abelianization $G/[G,G]$ is $\cong \mathbb{Z}_p^{d-1}$.

Theorem (Heisenberg Groups - cutoff)

Let $G := H_{p,d}$ with p prime and $d \geq 3$ fixed or diverging slowly and k be s.t. $1 \ll \log k \ll \log |G|$. Then, w.h.p. (as $p \to \infty$), the SRW on the G_k exhibits cutoff at time

$$t_\star(k,p,d) := \max\left\{\log_k n,\ t_0(k, p^{d-1})\right\}.$$

Heisenberg Matrix Groups

Let $G = H_{p,d}$ be the Heisenberg group of $d \times d$ uni-upper triangular matrices with integer entries mod p.

For $A, B \in G$ we have $(AB)_{i,i+1} = A_{i,i+1} + B_{i,i+1}$ and the Abelianization $G/[G, G]$ is $\cong \mathbb{Z}_p^{d-1}$.

Theorem (Heisenberg Groups - cutoff)

Let $G := H_{p,d}$ with p prime and $d \geq 3$ fixed or diverging slowly and k be s.t. $1 \ll \log k \ll \log |G|$. Then, w.h.p. (as $p \to \infty$), the SRW on the G_k exhibits cutoff at time

$$t_\star(k, p, d) := \max\big\{\log_k n,\ t_0(k, p^{d-1})\big\}.$$

Furthermore,

$$t_\star(k, p, d) \eqsim \begin{cases} t_0(k, p^{d-1}) & \textit{when} \quad k \leq (\log n)^{1+2/(d-2)}, \\ \log_k n & \textit{when} \quad k \geq (\log n)^{1+2/(d-2)}. \end{cases}$$

Lower bound

- For a probability μ and $X \sim \mu$, the **entropy** is

$$\mathrm{Ent}\mu := -\sum_{x} \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].$$

Lower bound

- For a probability μ and $X \sim \mu$, the **entropy** is

$$\mathrm{Ent}\mu := -\sum_{x} \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].$$

- As $-\log \mu(W(t)) = -\sum_{i=1}^{k} \log \nu(W_i(t))$,
  where μ and ν are the laws of $W(t)$ and $W_1(t)$, resp.,
  by CLT it is concentrated around its mean ($=$ entropy) when $k \gg 1$.

Lower bound

- For a probability μ and $X \sim \mu$, the **entropy** is
$$\mathrm{Ent}\mu := -\sum_{x} \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].$$
- As $-\log \mu(W(t)) = -\sum_{i=1}^{k} \log \nu(W_i(t))$,

  where μ and ν are the laws of $W(t)$ and $W_1(t)$, resp.,

  by CLT it is concentrated around its mean ($=$ entropy) when $k \gg 1$.
- We first argue that the walk is 'far from being mixed' (i.e. TV distance $1 - o(1)$) at time $t_- = t_0(1 - o(1))$, for some choice of $o(1)$ terms.
- A calculation shows that by changing t a little around t_0 we can change the entropy 'a lot',

Lower bound

- For a probability μ and $X \sim \mu$, the **entropy** is
$$\mathrm{Ent}\mu := -\sum_{x} \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].$$
- As $-\log \mu(W(t)) = -\sum_{i=1}^{k} \log \nu(W_i(t))$,

  where μ and ν are the laws of $W(t)$ and $W_1(t)$, resp.,

  by CLT it is concentrated around its mean ($=$ entropy) when $k \gg 1$.
- We first argue that the walk is 'far from being mixed' (i.e. TV distance $1 - o(1)$) at time $t_- = t_0(1 - o(1))$, for some choice of $o(1)$ terms.
- A calculation shows that by changing t a little around t_0 we can change the entropy 'a lot', i.e. by a diverging additive term, which is much larger than the typical fluctuations of $-\log \mu(W(t))$,

Lower bound

- For a probability μ and $X \sim \mu$, the **entropy** is

$$\mathrm{Ent}\mu := -\sum_{x} \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].$$

- As $-\log \mu(W(t)) = -\sum_{i=1}^{k} \log \nu(W_i(t))$,

  where μ and ν are the laws of $W(t)$ and $W_1(t)$, resp.,

  by CLT it is concentrated around its mean ($=$ entropy) when $k \gg 1$.

- We first argue that the walk is 'far from being mixed' (i.e. TV distance $1 - o(1)$) at time $t_- = t_0(1 - o(1))$, for some choice of $o(1)$ terms.

- A calculation shows that by changing t a little around t_0 we can change the entropy 'a lot', i.e. by a diverging additive term, which is much larger than the typical fluctuations of $-\log \mu(W(t))$, and that

$$\mathrm{Var}(\log \mu(W(t))) = (1 \pm o(1))\mathrm{Var}(\log \mu(W(t_0))).$$

Lower bound

- For a probability μ and $X \sim \mu$, the **entropy** is

$$\mathrm{Ent}\mu := -\sum_x \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].$$

- As $-\log \mu(W(t)) = -\sum_{i=1}^k \log \nu(W_i(t))$,

  where μ and ν are the laws of $W(t)$ and $W_1(t)$, resp.,

  by CLT it is concentrated around its mean ($=$ entropy) when $k \gg 1$.

- We first argue that the walk is 'far from being mixed' (i.e. TV distance $1-o(1)$) at time $t_- = t_0(1-o(1))$, for some choice of $o(1)$ terms.

- A calculation shows that by changing t a little around t_0 we can change the entropy 'a lot', i.e. by a diverging additive term, which is much larger than the typical fluctuations of $-\log \mu(W(t))$, and that

$$\mathrm{Var}(\log \mu(W(t))) = (1 \pm o(1))\mathrm{Var}(\log \mu(W(t_0))).$$

$\implies$ For some $t_- = (1-o(1))t_0$ and $\omega \gg 1$ w.h.p.
$\log n - \log \mu(W(t)) \geq 2\omega$ i.e. $\mu(W(t)) \geq e^{\omega}/n$.

Lower bound

- (CLT) $-\log \mu(W)$ is concentrated around its mean (entropy) when $k \gg 1$.
- A calculation shows that changing t a little around t_0 changed the entropy 'a lot', i.e. by much more than the typical fluctuations of $-\log \mu(W(t))$, and that $\mathrm{Var}(\log \mu(W(t))) = (1 \pm o(1))\mathrm{Var}(\log \mu(W(t_0)))$.
- $\implies$ For some $t_- = t = (1 - o(1))t_0$ and $\omega \gg 1$ w.h.p. $\mu(W(t)) \geq e^{\omega}/n$.
- On this event (which holds w.h.p.) $W(t)$ belongs to a set of size $\frac{n}{e^{\omega}} = o(n)$. (if all points in a set have probability at least p, its size is at most $1/p$).

Lower bound

- (CLT) $-\log \mu(W)$ is concentrated around its mean (entropy) when $k \gg 1$.
- A calculation shows that changing t a little around t_0 changed the entropy 'a lot', i.e. by much more than the typical fluctuations of $-\log \mu(W(t))$, and that $\mathrm{Var}(\log \mu(W(t))) = (1 \pm o(1))\mathrm{Var}(\log \mu(W(t_0)))$.
- $\implies$ For some $t_- = t = (1 - o(1))t_0$ and $\omega \gg 1$ w.h.p. $\mu(W(t)) \geq e^\omega/n$.
- On this event (which holds w.h.p.) $W(t)$ belongs to a set of size $\frac{n}{e^\omega} = o(n)$. (if all points in a set have probability at least p, its size is at most $1/p$).
- By projecting so does $S(t) = W(t) \cdot Z$, so for **all** Z not mixed. $\square$

Lower bound

- (CLT) $-\log \mu(W)$ is concentrated around its mean (entropy) when $k \gg 1$.
- A calculation shows that changing t a little around t_0 changed the entropy 'a lot', i.e. by much more than the typical fluctuations of $-\log \mu(W(t))$, and that $\mathrm{Var}(\log \mu(W(t))) = (1 \pm o(1))\mathrm{Var}(\log \mu(W(t_0)))$.
- $\implies$ For some $t_- = t = (1-o(1))t_0$ and $\omega \gg 1$ w.h.p. $\mu(W(t)) \geq e^{\omega}/n$.
- On this event (which holds w.h.p.) $W(t)$ belongs to a set of size $\frac{n}{e^{\omega}} = o(n)$. (if all points in a set have probability at least p, its size is at most $1/p$).
- By projecting so does $S(t) = W(t) \cdot Z$, so for **all** Z not mixed. $\square$
- Similarly, $t_m(1-o(1))$ can be shown to be a lower bound on the mixing time for **all** Z by considering $S(t)(mG)$, which is the induced walk on G/mG, and repeating the above argument to it.

Lower bound

- Similarly, $t_m(1 - o(1))$ can be shown to be a lower bound on the mixing time for **all** Z by considering $S(t)(mG)$, which is the induced walk on G/mG, and repeating the above argument to it:
- $S(t)$ is not mixed if this induced walk is not mixed.

Lower bound

- Similarly, $t_m(1 - o(1))$ can be shown to be a lower bound on the mixing time for **all** Z by considering $S(t)(mG)$, which is the induced walk on G/mG, and repeating the above argument to it:
- $S(t)$ is not mixed if this induced walk is not mixed.
- As before, by projection, $S(t)(mG)$ cannot be mixed if $(W(t) \mod m)$ w.h.p. belongs to a set of size $o(|G/mG|)$

Lower bound

- Similarly, $t_m(1-o(1))$ can be shown to be a lower bound on the mixing time for **all** Z by considering $S(t)(mG)$, which is the induced walk on G/mG, and repeating the above argument to it:
- $S(t)$ is not mixed if this induced walk is not mixed.
- As before, by projection, $S(t)(mG)$ cannot be mixed if $(W(t) \mod m)$ w.h.p. belongs to a set of size $o(|G/mG|)$; but as for $m=0$, this is the case for $t \leq (1-o(1))t_m$
  (by def of t_m, concentration of $\log(\mu_p(W(t) \mod m))$, with μ_p being the law of $W(t) \mod m$), and since we the mean of this r.v. can changes by a between time $(1-o(1))t_m$ and t_m by much more than the SD).

Upper bound: Warm up $G = \mathbb{Z}_p^d$ for p prime

- Let $t = (1 + o(1))t_p$. Write W for $W(t)$.
- We use our modified L_2 argument with $\mathrm{typ} = \{W \in \mathcal{W}\}$, where

$$\mathcal{W} = \{w \in \mathbb{Z}^k : \mathbb{P}[W \equiv w \mod p] \leq \delta\},$$

$$\text{where } \delta = \delta(n) = o(1/n) \text{ and } n := |G|.$$

- By the def. of t_p and concentration of $\log(\mu_p(W \mod p))$, where μ_p is the law of $W \mod p$, indeed $\mathbb{P}[\mathrm{typ}^c] = o(1)$ as desired, for some δ as above.

---

[6] If $\gcd(a, n) = 1$ then $g \mapsto g^a$ is invertible (by $g \mapsto g^b$ s.t. $ab \equiv 1 \mod n$) and so $X^a \sim \mathrm{Unif}(G)$ whenever $X \sim \mathrm{Unif}(G)$.

Upper bound: Warm up $G = \mathbb{Z}_p^d$ for p prime

- Let $t = (1 + o(1))t_p$. Write W for $W(t)$.
- We use our modified L_2 argument with $\mathrm{typ} = \{W \in \mathcal{W}\}$, where

$$\mathcal{W} = \{w \in \mathbb{Z}^k : \mathbb{P}[W \equiv w \bmod p] \leq \delta\},$$

$$\text{where } \delta = \delta(n) = o(1/n) \text{ and } n := |G|.$$

- By the def. of t_p and concentration of $\log(\mu_p(W \mod p))$, where μ_p is the law of $W \mod p$, indeed $\mathbb{P}[\mathrm{typ}^c] = o(1)$ as desired, for some δ as above.
- Recall $S = W \cdot Z = \sum_{i=1}^k W_i Z_i$ and $S' = W' \cdot Z$ with W' indep. of W.
- Note that if $W \not\equiv W' \bmod p$ then $S - S' \sim \mathrm{Uniform}(\mathbb{Z}_p^d)$.[6] Thus

$$n\mathbb{P}[S = S' \mid \mathrm{typ}] - 1 \leq n\mathbb{P}[W \equiv W' \bmod p \mid \mathrm{typ}]$$

[6] If $\gcd(a, n) = 1$ then $g \mapsto g^a$ is invertible (by $g \mapsto g^b$ s.t. $ab \equiv 1 \mod n$) and so $X^a \sim \mathrm{Unif}(G)$ whenever $X \sim \mathrm{Unif}(G)$.

Upper bound: Warm up $G = \mathbb{Z}_p^d$ for p prime

- Let $t = (1 + o(1))t_p$. Write W for $W(t)$.
- We use our modified L_2 argument with $\mathrm{typ} = \{W \in \mathcal{W}\}$, where

$$\mathcal{W} = \{w \in \mathbb{Z}^k : \mathbb{P}[W \equiv w \text{ mod } p] \leq \delta\},$$

$$\text{where } \delta = \delta(n) = o(1/n) \text{ and } n := |G|.$$

- By the def. of t_p and concentration of $\log(\mu_p(W \mod p))$, where μ_p is the law of $W \mod p$, indeed $\mathbb{P}[\mathrm{typ}^c] = o(1)$ as desired, for some δ as above.
- Recall $S = W \cdot Z = \sum_{i=1}^k W_i Z_i$ and $S' = W' \cdot Z$ with W' indep. of W.
- Note that if $W \not\equiv W' \text{ mod } p$ then $S - S' \sim \mathrm{Uniform}(\mathbb{Z}_p^d)$.[6] Thus

$$n\mathbb{P}[S = S' \mid \mathrm{typ}] - 1 \leq n\mathbb{P}[W \equiv W' \text{ mod } p \mid \mathrm{typ}] \lesssim n\delta = o(1). \quad \square$$

[6] If $\gcd(a, n) = 1$ then $g \mapsto g^a$ is invertible (by $g \mapsto g^b$ s.t. $ab \equiv 1 \mod n$) and so $X^a \sim \mathrm{Unif}(G)$ whenever $X \sim \mathrm{Unif}(G)$.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- The idea extends to general G. Want $\mathcal{W} \subset \mathbb{Z}^k$ such that $W \in \mathcal{W}$ w.h.p. and

$$|G|\rho - 1 = o(1), \text{ where } \rho := \mathbb{P}[S = S' \mid W, W' \in \mathcal{W}].$$

$$\rho = \underbrace{\mathbb{P}[W = W' \mid W, W' \in \mathcal{W}]}_{=A} + \underbrace{\mathbb{P}[S = S' \mid W \neq W' \in \mathcal{W}]}_{B}.$$

[7] By concentration of $\log(\mu(W))$, the def. of t_0 and the rapid change of entropy.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- The idea extends to general G. Want $\mathcal{W} \subset \mathbb{Z}^k$ such that $W \in \mathcal{W}$ w.h.p. and

$$|G|\rho - 1 = o(1), \text{ where } \rho := \mathbb{P}[S = S' \mid W, W' \in \mathcal{W}].$$

$$\rho = \underbrace{\mathbb{P}[W = W' \mid W, W' \in \mathcal{W}]}_{=A} + \underbrace{\mathbb{P}[S = S' \mid W \neq W' \in \mathcal{W}]}_{B}.$$

If $t \geq (1 + o(1))t_0$ and $\mathcal{W} \subset \mathcal{W}_0 := \{w \in \mathbb{Z}^k : \mathbb{P}[W = w] \leq \varepsilon_0/n\}$, for some appropriate choice of $\varepsilon_0 = o(1)$, then as before: $nA \lesssim \varepsilon_0 = o(1)$, and we will have $W \in \mathcal{W}_0$ w.h.p.[7]

[7] By concentration of $\log(\mu(W))$, the def. of t_0 and the rapid change of entropy.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- The idea extends to general G. Want $\mathcal{W} \subset \mathbb{Z}^k$ such that $W \in \mathcal{W}$ w.h.p. and

$$|G|\rho - 1 = o(1), \text{ where } \rho := \mathbb{P}[S = S' \mid W, W' \in \mathcal{W}].$$

$$\rho = \underbrace{\mathbb{P}[W = W' \mid W, W' \in \mathcal{W}]}_{=A} + \underbrace{\mathbb{P}[S = S' \mid W \neq W' \in \mathcal{W}]}_{B}.$$

If $t \geq (1 + o(1))t_0$ and $\mathcal{W} \subset \mathcal{W}_0 := \{w \in \mathbb{Z}^k : \mathbb{P}[W = w] \leq \varepsilon_0/n\}$, for some appropriate choice of $\varepsilon_0 = o(1)$, then as before:
$nA \lesssim \varepsilon_0 = o(1)$, and we will have $W \in \mathcal{W}_0$ w.h.p.[7]

Given $(W, W') = (w, w)$ we have $S - S' \sim \mathrm{Unif}(\mathfrak{g}G)$, where $\mathfrak{g} := \gcd(V_1, \ldots, V_k, n)$ (recall $V_i := W_i - W_i'$),

$$\implies |G|\mathbb{P}[S = S' \mid W, W', \mathfrak{g}] = \frac{|G|}{|\mathfrak{g}G|} \leq \mathfrak{g}^d \wedge n.$$

[7] By concentration of $\log(\mu(W))$, the def. of t_0 and the rapid change of entropy.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- The idea extends to general G. Want $\mathcal{W} \subset \mathbb{Z}^k$ such that $W \in \mathcal{W}$ w.h.p. and

$$|G|\rho - 1 = o(1), \text{ where } \rho := \mathbb{P}[S = S' \mid W, W' \in \mathcal{W}].$$

$$\rho = \underbrace{\mathbb{P}[W = W' \mid W, W' \in \mathcal{W}]}_{=A} + \underbrace{\mathbb{P}[S = S' \mid W \neq W' \in \mathcal{W}]}_{B}.$$

If $t \geq (1 + o(1))t_0$ and $\mathcal{W} \subset \mathcal{W}_0 := \{w \in \mathbb{Z}^k : \mathbb{P}[W = w] \leq \varepsilon_0/n\}$, for some appropriate choice of $\varepsilon_0 = o(1)$, then as before: $nA \lesssim \varepsilon_0 = o(1)$, and we will have $W \in \mathcal{W}_0$ w.h.p.[7]

Given $(W, W') = (w, w)$ we have $S - S' \sim \mathrm{Unif}(\mathfrak{g}G)$, where $\mathfrak{g} := \gcd(V_1, \ldots, V_k, n)$ (recall $V_i := W_i - W_i'$),

$$\Longrightarrow \ |G|\mathbb{P}[S = S' \mid W, W', \mathfrak{g}] = \frac{|G|}{|\mathfrak{g}G|} \leq \mathfrak{g}^d \wedge n.$$

$$\Longrightarrow \ |G|B - 1 \leq \mathbb{E}[(\mathfrak{g}^d \wedge n)\mathbf{1}\{\mathfrak{g} > 1\} \mid W \neq W' \in \mathcal{W}].$$

[7] By concentration of $\log(\mu(W))$, the def. of t_0 and the rapid change of entropy.

Upper bound for general abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Our problem is reduced to arguing that for $t = (1 + o(1))T$ for some choice of $\mathcal{W} \supset \mathcal{W}_0$ we have

$$(*) := \mathbb{E}[(\mathfrak{g}^d \wedge n)\mathbf{1}\{\mathfrak{g} > 1\} \mid W \neq W' \in \mathcal{W}] = o(1).$$

Upper bound for general abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Our problem is reduced to arguing that for $t = (1 + o(1))T$ for some choice of $\mathcal{W} \supset \mathcal{W}_0$ we have

$$(*) := \mathbb{E}[(\mathfrak{g}^d \wedge n)\mathbf{1}\{\mathfrak{g} > 1\} \mid W \neq W' \in \mathcal{W}] = o(1).$$

- A calculation reveals that $T \lesssim kn^{2/k} \log k$ and thus w.h.p. $\max_i |W_i| \leq r_* = n^{1/k}(\log k)^2$.

Upper bound for general abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Our problem is reduced to arguing that for $t = (1 + o(1))T$ for some choice of $\mathcal{W} \supset \mathcal{W}_0$ we have

$$(*) := \mathbb{E}[(\mathfrak{g}^d \wedge n)\mathbf{1}\{\mathfrak{g} > 1\} \mid W \neq W' \in \mathcal{W}] = o(1).$$

- A calculation reveals that $T \lesssim kn^{2/k} \log k$ and thus w.h.p. $\max_i |W_i| \leq r_* = n^{1/k}(\log k)^2$.
- Including this constraint in $\mathcal{W}$, we only need to consider $\mathfrak{g} \in [2, 2r_*]$.

Upper bound for general abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Our problem is reduced to arguing that for $t = (1 + o(1))T$ for some choice of $\mathcal{W} \supset \mathcal{W}_0$ we have

$$(*) := \mathbb{E}[(\mathfrak{g}^d \wedge n)\mathbf{1}\{\mathfrak{g} > 1\} \mid W \neq W' \in \mathcal{W}] = o(1).$$

- A calculation reveals that $T \lesssim kn^{2/k} \log k$ and thus w.h.p. $\max_i |W_i| \leq r_* = n^{1/k}(\log k)^2$.
- Including this constraint in $\mathcal{W}$, we only need to consider $\mathfrak{g} \in [2, 2r_*]$. Using $\mathbb{P}[D \mid E] \leq P[D]/P[E] = (1 + o(1))P[E]$ if $P[E^c] = o(1)$:

$$(*) \leq (1 + o(1)) \sum_{\ell=2}^{2r_*} (\ell^d \wedge n)\mathbb{P}[\ell \text{ divides all of } V_1, \ldots, V_k].$$

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Recall $r_* := n^{1/k}(\log k)^2$. Want

$$(**) := \sum_{\ell=2}^{2r_*} (\ell^d \wedge n)\mathbb{P}[\ell \text{ divides all of } V_1, \dots, V_k] = o(1).$$

- We will show

$$\mathbb{P}[\ell \text{ divides } V_1] \le \mathbb{P}[V_1 = 0] + 1/\ell \approx \frac{C'}{n^{1/k}} + 1/\ell.$$

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Recall $r_* := n^{1/k}(\log k)^2$. Want

$$(**) := \sum_{\ell=2}^{2r_*} (\ell^d \wedge n)\mathbb{P}[\ell \text{ divides all of } V_1, \ldots, V_k] = o(1).$$

- We will show

$$\mathbb{P}[\ell \text{ divides } V_1] \leq \mathbb{P}[V_1 = 0] + 1/\ell \approx \frac{C'}{n^{1/k}} + 1/\ell.$$

- Substituting in $(**)$ and using the fact that $V_1, \ldots, V_k$ are i.i.d.,

$$(**) \leq \sum_{\ell=2}^{2r_*} (\ell^d \wedge n)[\frac{C'}{n^{1/k}} + 1/\ell]^k = o(1),$$

if $k \ll \log n$ and k is "a bit" larger than d.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

- Recall $r_* := n^{1/k}(\log k)^2$. Want

$$(**) := \sum_{\ell=2}^{2r_*} (\ell^d \wedge n)\mathbb{P}[\ell \text{ divides all of } V_1, \ldots, V_k] = o(1).$$

- We will show

$$\mathbb{P}[\ell \text{ divides } V_1] \le \mathbb{P}[V_1 = 0] + 1/\ell \approx \frac{C'}{n^{1/k}} + 1/\ell.$$

- Substituting in $(**)$ and using the fact that $V_1, \ldots, V_k$ are i.i.d.,

$$(**) \le \sum_{\ell=2}^{2r_*} (\ell^d \wedge n)[\frac{C'}{n^{1/k}} + 1/\ell]^k = o(1),$$

if $k \ll \log n$ and k is “a bit” larger than d. How much is “a bit” depends on k. For $k \ll \sqrt{\log n / \log\log\log n}$ it turns out $k - d \gg 1$ suffices.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

Proof of: $\mathbb{P}[\ell \text{ divides } V_1] \leq \mathbb{P}[V_1 = 0] + 1/\ell$

- Given $V_1 \neq 0$ (by unimodality) the conditional law of $|V_1|$ is unimodal, and thus can be written as a mixture of uniform distributions $\mathrm{Unif}(\{1, \ldots, Y\})$, where Y is random.

General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

Proof of: $\mathbb{P}[\ell \text{ divides } V_1] \leq \mathbb{P}[V_1 = 0] + 1/\ell$

- Given $V_1 \neq 0$ (by unimodality) the conditional law of $|V_1|$ is unimodal, and thus can be written as a mixture of uniform distributions $\mathrm{Unif}(\{1, \ldots, Y\})$, where Y is random.

  $\Longrightarrow$ the probability that ℓ divides V_1, given $V_1 \neq 0$, is at most $\frac{1}{\ell}$. $\square$

General abelian G of size n

To treat $k \asymp \log n$ and to allow $k-d$ to diverge arbitrary slowly for $k \gg \sqrt{\log n}$ we can no longer use the bound $\frac{|G|}{|\ell G|} \leq \ell^d \wedge n$ and instead need to show for $t=(1+\delta)T$ that for some "typical" $\mathcal{W}$,

$$(***) := \sum_{\ell \in [2,2r_*]:\ \ell | n} \frac{|G|}{|\ell G|} \mathbb{P}[\ell \text{ divides all of } V_1, \dots, V_k \mid W, W' \in \mathcal{W}] = o(1).$$

General abelian G of size n

Need to show for $t = (1+\delta)T$ that for some 'typical' $\mathcal{W}$,

$$(***) := \sum_{\ell \in [2,2r_*]:\ \ell | n} \frac{|G|}{|\ell G|} \mathbb{P}[W \equiv W' \mod \ell \mid W, W' \in \mathcal{W}] = o(1).$$

- For some $\varepsilon = o(1/r_*)$ consider

$$\mathcal{W}_\ell := \{w \in \mathbb{Z}^k : \mathbb{P}[W = w \mod a] \leq \varepsilon / |G/\ell G|\},$$

and $\mathcal{W} := \mathcal{W}_0 \cap \{w : \max_i |w_i| \leq r_*\} \cap (\bigcap_{[2,2r_*]:\ \ell | n} \mathcal{W}_\ell)$.

General abelian G of size n

Need to show for $t = (1+\delta)T$ that for some 'typical' $\mathcal{W}$,

$$(***) := \sum_{\ell \in [2,2r_*]:\ \ell | n} \frac{|G|}{|\ell G|} \mathbb{P}[W \equiv W' \mod \ell \mid W, W' \in \mathcal{W}] = o(1).$$

- For some $\varepsilon = o(1/r_*)$ consider

$$\mathcal{W}_\ell := \{w \in \mathbb{Z}^k : \mathbb{P}[W = w \mod a] \le \varepsilon/|G/\ell G|\},$$

  and $\mathcal{W} := \mathcal{W}_0 \cap \{w : \max_i |w_i| \le r_*\} \cap (\bigcap_{[2,2r_*]:\ \ell|n} \mathcal{W}_\ell)$.
- For this $\mathcal{W}$: $(***) \le \varepsilon r_* = o(1)$.
- Turns out we can take $\varepsilon = e^{-\Omega(\delta)k}$ and still have that $\mathcal{W}$ is 'typical', and for $k \gg \sqrt{\log n}$ can pick $\delta = o(1)$ so that indeed $e^{-\Omega(\delta)k} \ll n^{-1/k}(\log k)^{-2} \le 1/r_*$. $\square$

Thank you for your attention.