Random Cayley Graphs

Jonathan Hermon, Sam Thomas

University of British Columbia

Stanford University, 11th May 2020
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 $Z := [Z_1, \ldots, Z_k] \subseteq G$ is the graph with vertex set $G$ and edge set

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

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

---

1 We sample with replacements. Since $k \ll \sqrt{n}$ we can also sample without. $\ll$ means little $o$ and $\lesssim$ means big $O$. 
The **Cayley digraph** of a group $G$ with respect to $Z := [Z_1, \ldots, Z_k] \subseteq G$ is the graph with vertex set $G$ and edge set

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

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

Throughout, we choose $k$ random elements $Z_1, \ldots, Z_k$ 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 $Z_1, \ldots, Z_k$ **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 $Z := [Z_1, ..., Z_k] \subseteq G$ is the graph with vertex set $G$ and edge set

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

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

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 $Z$ 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

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

In 1986 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 (1992) for $k$ s.t. $\log k \gg \log \log |G|$: cutoff at time $\log_k |G|$.
- Dou & Hildebrand (1994) - cutoff for abelian $G$ when $k = \lceil (\log n)^a \rceil$ for $a > 1$ at time $\log_k / \log n$ (where $n := |G|$).
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

\[
\Phi := \min_{A \subset G: |A| \leq |G|/2} \frac{|E(A, A^c)|}{2k|A|} \geq 1 - \varepsilon
\]

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

\[
\Phi := \min_{A \subset G: |A| \leq |G|/2} \frac{|E(A, A^c)|}{2k|A|} \geq 1 - \varepsilon
\]

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

- By Cheeger’s ineq. \( \Phi \) 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

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

  \( (k \geq (1 + \delta)d(G) \) 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?

Here $G_{ab} = G/\langle [G,G] \rangle$ is the abelianization of $G$, and $\langle [G,G] \rangle$ is its commutator (the group generated by $\{ghg^{-1}h^{-1} : g, h \in G\}$).

- Open Problem: For $G = S_n$ 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 $n^3 \log n$.)
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 $k \geq C \log |G^{ab}|$?

Here $G^{ab} := G/[G, G]$ 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 $G = S_n$ 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 $n^3 \log n$.)
Shapira and Zuck (18) (improving Marklof and Strömbergsson) - For a sequence of abelian $G^{(n)}$ with fixed $d(G^{(n)}) = d$ and fixed $k \geq d$:

\[(\ast) \quad |G^{(n)}|^{-1/k} \text{Diameter}(G_k^{(n)})\]

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 $(\ast)$ as $p \to \infty$, with $|G|$ replaced with $|G^{ab}|$. 
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 $\text{Diameter}(G_k) = R + o(R)$. 
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 $\text{Diameter}(G_k) = R + o(R)$.

- H. & Thomas (19) - similar results hold for the Heisenberg group with $|G|$ above replaced with $|G^{ab}|$. 
The total variation (TV) distance of two distributions $\mu$ and $\nu$ on the same finite set $G$ is

$$\|\mu - \nu\|_{TV} := \frac{1}{2} \sum_{x \in G} |\mu(x) - \nu(x)|.$$ 

The TV $\varepsilon$-mixing time of a Markov chain $(X_t)_{t \geq 0}$ is

$$t_{\text{mix}}(\varepsilon) := \inf \left\{ t : \max_x \|P^t_x[X_t = \cdot] - \pi\|_{TV} \leq \varepsilon \right\},$$

where $\pi$ is the stationary distribution.
The total variation (TV) distance is \( \|\mu - \nu\|_{TV} := \frac{1}{2} \sum_x |\mu(x) - \nu(x)| \).

The TV \( \epsilon \)-mixing time is \( t_{mix}(\epsilon) := \inf \{ t : \max_x \|\mathbb{P}_x[X_t = \cdot] - \pi\|_{TV} \leq \epsilon \} \), where \( \pi \) is the stationary distribution.

A sequence of Markov chains exhibits (TV) **cutoff** if the \( \epsilon \)-mixing time is asymptotically indep. of \( \epsilon \):

\[
\lim_{n \to \infty} t_{mix}^{(n)}(\epsilon)/t_{mix}^{(n)}(1 - \epsilon) = 1, \quad \forall 0 < \epsilon < 1
\]

(1)

(the superscript \( (n) \) indicates that this is the mixing time of the \( n \)th 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 $\varepsilon$-mixing time is asymptotically indep. of $\varepsilon$:

$$\lim_{n \to \infty} \frac{t_{\text{mix}}^{(n)}(\varepsilon)}{t_{\text{mix}}^{(n)}(1 - \varepsilon)} = 1, \quad \forall 0 < \varepsilon < 1.$$  

(2)

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

**Figure:** The worst case TV distance of the $n$th chain, $d_n(t) := \max_x \| \mathbb{P}_x [X_t = \cdot] - \pi \|_{TV}$ 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 $\geq 3$.
- 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 $\geq 3$) the edges of a random perfect matching of the vertices.

---

$^2$A random $n$ lift of $(V, E)$ is generated by taking $n$ copies $v_1, \ldots, v_n$ of each $v \in V$ and for each $uv \in E$ connect $u_i$ with $v_{\tau_{uv}(i)}$, where $\tau_{uv}$ 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\(^4\) $\mathbb{Z}^k_m$ 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 $\log |G^{ab}|$.

\(^3\)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 $\log |G^{ab}|$. 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 $\mu$ on $G$ the $L_2$ distance from the uniform distribution $\pi$ is defined as

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

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

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 $Z := [Z_1, \ldots, Z_k]$; i.e., given the graph) then

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

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

- 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 $Z := [Z_1, \ldots, Z_k]$; i.e., given the graph) then

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

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

- 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 $Z := [Z_1, \ldots, Z_k]$, then

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

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

\[
\mathbb{E}[\|\mathbb{P}[S(t) = \cdot | Z] - \pi\|_2^2] = |G|\mathbb{P}[S(t) = S'(t)] - 1.
\]

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 $o(1)$. 

Jonathan Hermon (UBC)  
Random Cayley Graphs  
Stanford University, 2020  
21 / 41
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

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

Crucially, the r.h.s. is an \textbf{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 $Z := [Z_1, \ldots, Z_k]$, then

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

- 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 $Z := [Z_1, \ldots, Z_k]$, then
  \[ \mathbb{E}[\|P[S(t) = \cdot | Z] - \pi\|^2_{2,\pi}] = |G|P[S(t) = S'(t)] - 1. \]

- 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 $S^{-1}S' \sim \text{Unif}(G) \implies P[S(t) = S'(t) | \text{this event}] = 1/|G|.$

  (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:

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

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

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

(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:

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

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

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

(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:

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

$$\leq \mathbb{P}[W(t) \notin \text{typ}] + \mathbb{E}[\| \mathbb{P}[S(t) = \cdot \mid Z, W(t) \in \text{typ}] - \pi \|_{TV}]$$
A modified $L_2$ bound

We use the modified $L_2$ bound:

$$
\frac{1}{2} \mathbb{E}[\|\mathbb{P}[S(t) = \cdot | Z] - \pi\|_{TV}]^2
\leq |G| \mathbb{P}[S(t) = S'(t) | W(t), W'(t) \in \text{typ}] - 1 + \mathbb{P}[W(t) \notin \text{typ}],
$$

where $\text{typ}$ is a certain ‘typical’ event (i.e. $\mathbb{P}[W(t) \notin \text{typ}] = o(1)$)
(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 | Z] - \pi\|_{TV}]
\leq \mathbb{P}[W(t) \notin \text{typ}] + \mathbb{E}[\|\mathbb{P}[S(t) = \cdot | Z, W(t) \in \text{typ}] - \pi\|_{TV}]
\leq \mathbb{E}[\|\mathbb{P}[S(t) = \cdot | W(t) \in \text{typ}, Z] - \pi\|_{TV}]^2
\quad \text{(Cauchy-Schwarz)}
\leq \mathbb{E}[\|\mathbb{P}[S(t) = \cdot | W(t) \in \text{typ}, Z] - \pi\|_{2, \pi}]^2
= |G| \mathbb{P}[S(t) = S'(t) | W(t), W'(t) \in \text{typ}] - 1.
\square
Warm up - a proof of Dou’s result: For $\log k \gg \log \log n$ cutoff at time $\log_2 k n$ (where $n := |G|$)

Proof: $t := \log_2 k n$ 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 $\text{typ}$

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

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

We want a smart choice of $\text{typ}$ 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 \( \text{typ} \)

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

\( \rho := \mathbb{P}[\text{no generator picked once in } W \text{ and } 0 \text{ in } W' \text{ or vice-versa} \mid W, W' \in \text{typ}] \). We want a smart choice of \( \text{typ} \) so that \( \rho := o(1/n) \).

We take \( \text{typ} \) 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 \( t = \log_{\log k / \log n} n \).
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

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

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

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 \( t = \log_k \log n \approx \log_{2k} n \) 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 \( \text{typ} \)

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

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

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

We take \( \text{typ} \) 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 \( \text{typ} \) 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 $\text{typ}$ to be the event that roughly the right number of generators are picked once and zero times.

---

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 $\text{typ}$ to be the event that roughly the right number of generators are picked once and zero times.

Then $\mathbb{P}[W \notin \text{typ}] \leq n^{-2c}$ 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 $2n^{-2c}$,

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

(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 \( \text{typ} \) to be the event that roughly the right number of generators are picked once and zero times.

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

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

\[
\implies t_{\text{mix}}(1/n^c) \leq C' \log n
\]

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

Conclude using

\[
\frac{1}{\text{gap}} \leq \frac{t_{\text{mix}}(\delta/2)}{\log(1/\delta)}
\]

\(^5\)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 \( \text{typ} \) to be the event that roughly the right number of generators are picked once and zero times.

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

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

\[ \Rightarrow \quad t_{\text{mix}}(1/n^c) \leq C' \log n \]

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

Conclude using

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

---

\( \text{5The choice of } C' \text{ depends on } C. \text{ The larger } C \text{ is, the smaller } C' \text{ is.} \)
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. \( \pm 1 \) 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. $\pm 1$ 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. $\pm 1$ with equal probability. ‘Coupon collector for groups’:

Step 1 (inspired by Erdős & 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. \( \pm 1 \) 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 $\log |G/mG|$, where for $m \in \mathbb{N}$, $mG := \{mg : g \in G\}$, (If $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$, then $mG \cong \mathbb{Z}_{m_1/\gcd(m_1,m)} \oplus \cdots \oplus \mathbb{Z}_{m_d/\gcd(m_d,m)}$)

- 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 $\mu$ 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

\[
t_*(k, p, d) := \max\{\log_k n, \ t_0(k, p^{d-1})\}.
\]
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

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

Furthermore,

$$t_*(k, p, d) \sim \begin{cases} 
  t_0(k, p^{d-1}) & \text{when } k \leq (\log n)^{1+2/(d-2)}, \\
  \log_k n & \text{when } k \geq (\log n)^{1+2/(d-2)}. 
\end{cases}$$
For a probability $\mu$ and $X \sim \mu$, the entropy is

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

- For a probability $\mu$ and $X \sim \mu$, the entropy is
  \[
  \text{Ent}\mu := -\sum_x \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].
  \]

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

Jonathan Hermon (UBC)  Random Cayley Graphs  Stanford University, 2020  31 / 41
Lower bound

- For a probability $\mu$ and $X \sim \mu$, the **entropy** is
  \[
  \text{Ent}\mu := - \sum_x \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].
  \]

- As $-\log \mu(W(t)) = - \sum_{i=1}^k \log \nu(W_i(t))$, where $\mu$ and $\nu$ 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 $t_\sim = t_0(1 - o(1))$, 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 $\mu$ and $X \sim \mu$, the **entropy** is
  \[
  \text{Ent}\mu := - \sum_x \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)].
  \]

- As $-\log \mu(W(t)) = - \sum_{i=1}^k \log \nu(W_i(t))$, where $\mu$ and $\nu$ 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 $\mu$ and $X \sim \mu$, the entropy is
  \[ \text{Ent}\mu := -\sum_x \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)]. \]

- As $-\log \mu(W(t)) = -\sum_{i=1}^k \log \nu(W_i(t))$, where $\mu$ and $\nu$ 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
  \[ \text{Var}(\log \mu(W(t))) = (1 \pm o(1))\text{Var}(\log \mu(W(t_0))). \]
Lower bound

- For a probability $\mu$ and $X \sim \mu$, the entropy is
  \[ \text{Ent}_\mu := -\sum_x \mu(x) \log \mu(x) = -\mathbb{E}[\log \mu(X)]. \]

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

- We first argue that the walk is ‘far from being mixed’ (i.e. TV distance $1 - o(1)$) at time $t_\sim = t_0(1 - o(1))$, 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
  \[ \text{Var}(\log \mu(W(t))) = (1 \pm o(1))\text{Var}(\log \mu(W(t_0))). \]

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

- $\implies$ For some $t_\pm = t = (1 - o(1))t_0$ 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 $\text{Var}(\log \mu(W(t))) = (1 \pm o(1)) \text{Var}(\log \mu(W(t_0)))$.
- $\implies$ 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.
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 $\text{Var}(\log \mu(W(t))) = (1 \pm o(1))\text{Var}(\log \mu(W(t_0)))$.
- $\implies$ For some $t_\approx = t = (1 - o(1))t_0$ 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.
- 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.
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 $o(|G/mG|)$.
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 $\mu_p$ 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 $\text{typ} = \{W \in \mathcal{W}\}$, where

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

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

---

6If $\gcd(a, n) = 1$ then $g \mapsto g^a$ is invertible (by $g \mapsto g^b$ s.t. $ab \equiv 1 \mod n$) and so $X^a \sim \text{Unif}(G)$ whenever $X \sim \text{Unif}(G)$. 
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 typ = $\{W \in \mathcal{W}\}$, where
  \[
  \mathcal{W} = \{w \in \mathbb{Z}^k : \mathbb{P}[W \equiv w \text{ mod } p] \leq \delta\},
  \]
  where $\delta = \delta(n) = o(1/n)$ and $n := |G|$.
- By the def. of $t_p$ and concentration of $\log(\mu_p(W \text{ mod } p))$, where $\mu_p$ is the law of $W \text{ mod } p$, indeed $\mathbb{P}[^{\text{typ}}\mathcal{C}] = o(1)$ as desired, for some $\delta$ as above.
- Recall $S = W \cdot Z = \sum_{i=1}^k W_iZ_i$ and $S' = W' \cdot Z$ with $W'$ indep. of $W$.
- Note that if $W \not\equiv W' \text{ mod } p$ then $S - S' \sim \text{Uniform}(\mathbb{Z}_p^d)$.
  Thus
  \[
  n\mathbb{P}[S = S' \mid \text{typ}] - 1 \leq n\mathbb{P}[W \equiv W' \text{ mod } p \mid \text{typ}]
  \]

---

If $\gcd(a, n) = 1$ then $g \mapsto g^a$ is invertible (by $g \mapsto g^b$ s.t. $ab \equiv 1 \text{ mod } n$) and so $X^a \sim \text{Unif}(G)$ whenever $X \sim \text{Unif}(G)$.
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 $\text{typ} = \{ W \in \mathcal{W} \}$, where

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

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

- By the def. of $t_p$ and concentration of $\log(\mu_p(W \mod p))$, where $\mu_p$ is the law of $W \mod p$, indeed $\mathbb{P}[\text{typ}^c] = o(1)$ as desired, for some $\delta$ 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 $W \not\equiv W' \mod p$ then $S - S' \sim \text{Uniform}(\mathbb{Z}_p^d)$.\(^6\) Thus

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

---

\(^6\)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 $X^a \sim \text{Unif}(G)$ whenever $X \sim \text{Unif}(G)$. 

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

$$|G|\rho - 1 = o(1),$$

where $\rho := \mathbb{P}[S = S' \mid W, W' \in \mathcal{W}]$.

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

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

\[
\rho = \underbrace{\mathbb{P}[W = W' \mid W, W' \in \mathcal{W}]}_{=A} + \underbrace{\mathbb{P}[S = S' \mid W \not= W' \in \mathcal{W}]}_{=B}.
\]

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 \( \mathcal{W} \in \mathcal{W}_0 \) 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

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

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

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 $S - S' \sim \text{Unif}(gG)$, where $g := \gcd(V_1, \ldots, V_k, n)$ (recall $V_i := W_i - W'_i$),

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

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

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

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:

$nA \leq \varepsilon_0 = o(1)$, and we will have $W \in \mathcal{W}_0$ w.h.p.

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

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

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

---

7By 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

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

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

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

Jonathan Hermon (UBC)
Random Cayley Graphs
Stanford University, 2020
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

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

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

Including this constraint in \( \mathcal{W} \), we only need to consider \( g \in [2, 2r_*] \).
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

\[
(\ast) := \mathbb{E}[(\ell^d \land n) \mathbf{1}\{\ell > 1\} \mid \mathcal{W} \neq \mathcal{W}' \in \mathcal{W}] = o(1).
\]

- A calculation reveals that $T \lesssim kn^{2/k} \log k$ and thus w.h.p.

\[
\max_i |\mathcal{W}_i| \leq r_* = n^{1/k}(\log k)^2.
\]

- Including this constraint in $\mathcal{W}$, we only need to consider $\ell \in [2, 2r_*]$.

Using $\mathbb{P}[D \mid E] \leq P[D]/P[E] = (1 + o(1))P[E]$ if $P[E^c] = o(1)$:

\[
(\ast) \leq (1 + o(1)) \sum_{\ell=2}^{2r_*} (\ell^d \land n) \mathbb{P}[\ell \text{ divides all of } V_1, \ldots, V_k].
\]
General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

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

- We will show
  
  $$ \mathbb{P}[\ell \text{ divides } V_1] \leq \mathbb{P}[V_1 = 0] + 1/\ell \approx \frac{C'}{n^{1/k}} + 1/\ell. $$
General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

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

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

- We will show

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

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

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

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

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

- 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.,

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

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 $\text{Unif}([1, \ldots, Y])$, where $Y$ is random.
General abelian $G = \mathbb{Z}_{m_1} \oplus \cdots \oplus \mathbb{Z}_{m_d}$

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

- 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 $\operatorname{Unif}\{1, \ldots, Y\}$, where $Y$ is random.

$\implies$ the probability that $\ell$ divides $V_1$, given $V_1 \neq 0$, is at most $1/\ell$.  \qed
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 $\frac{|G|}{|\ell G|} \leq \ell^d \land n$ and instead need to show for $t = (1 + \delta)T$ that for some ”typical” $\mathcal{W}$,

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

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

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

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

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

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’ $\mathcal{W}$,

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

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

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

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

- For this $\mathcal{W}$: $(**\star\star\star) \leq \varepsilon r_* = o(1)$.

- 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 $e^{-\Omega(\delta)k} \ll n^{-1/k}(\log k)^{-2} \leq 1/r_*$. 

\[\square\]
Thank you for your attention.