Cutoff for the Asymmetric Riffle Shuffle

Mark Sellke (Stanford)

University of Chicago Probability Seminar

January 14, 2022

- How many shuffles are needed to mix a deck of cards?

Card Shuffling

- How many shuffles are needed to mix a deck of cards?

- Depends on how you shuffle...

- How many shuffles are needed to mix a deck of cards?

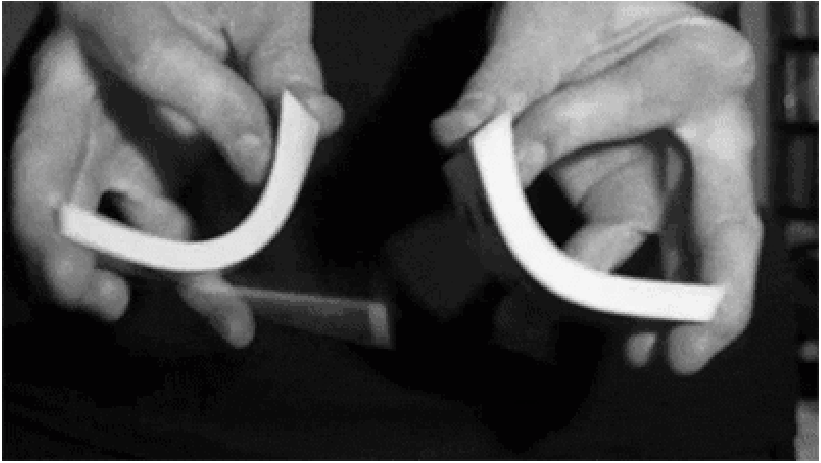
- Depends on how you shuffle...
  - Cut repeatedly
  - Top to random
  - Random to random
  - ...

Card Shuffling

- How many shuffles are needed to mix a deck of cards?

- Depends on how you shuffle...
    - Cut repeatedly
    - Top to random
    - Random to random
    - ...
- Today: **riffle shuffle**

Just So We're Clear...



The GSR Shuffle

- Gilbert-Shannon-Reeds (GSR) model for riffle shuffles:
  1. Cut deck into $Bin(N, 1/2)$ sized piles.
  2. Riffle the piles together uniformly at random.

- Gilbert-Shannon-Reeds (GSR) model for riffle shuffles:
  1. Cut deck into $Bin(N, 1/2)$ sized piles.
  2. Riffle the piles together uniformly at random.
- Equivalent to 2: if current pile sizes are A and B, drop next card from first pile with probability $\frac{A}{A+B}$.

- Gilbert-Shannon-Reeds (GSR) model for riffle shuffles:
  1. Cut deck into $Bin(N, 1/2)$ sized piles.
  2. Riffle the piles together uniformly at random.
- Equivalent to 2: if current pile sizes are A and B, drop next card from first pile with probability $\frac{A}{A+B}$.
- [Aldous 83, Bayer-Diaconis 92]: total variation mixing occurs after

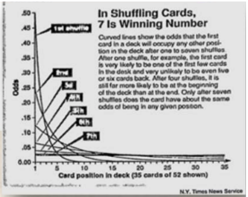
$$t_{\mathrm{mix}} = \frac{3\log_2(N)}{2} \pm O(1)$$

GSR shuffles.
- Reminder: for probability distributions P, Q on finite set X, total variation distance is

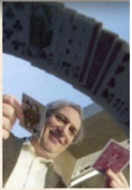
$$d_{TV}(P, Q) = \frac{1}{2}\sum_{x \in X} |P(x) - Q(x)|.$$

‘7 shuffles suffice’

The New York Times January 9, 1990



It takes just seven ordinary, imperfect shuffles to mix a deck of cards thoroughly, researchers have found. Fewer are not enough and more do not significantly improve the mixing.



The mathematical proof, discovered after studies of results from elaborate computer calculations and careful observation of card games, confirms the intuition of many gamblers, bridge enthusiasts and casual players that most shuffling is inadequate.

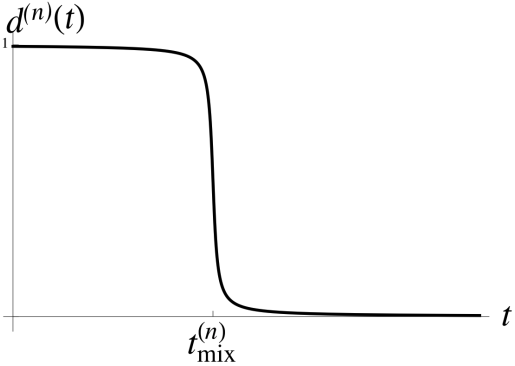
...

By saying that the deck is completely mixed after seven shuffles, Dr. Diaconis and Dr. Bayer mean that every arrangement of the 52 cards is equally likely or that any card is as likely to be in one place as in another.

The cards do get more and more randomly mixed if a person keeps on shuffling more than seven times, but seven shuffles is a transition point, the first time that randomness is close. Additional shuffles do not appreciably alter things...

(Arrangement by Eyal Lubetzky)

- Cutoff phenomenon: sharp transition in distance to stationarity at *mixing time* $t_{\mathrm{mix}} \pm o(t_{\mathrm{mix}})$. Many examples:
  - Transpositions on S_n (Diaconis-Shahshahani)
  - Glauber dynamics for high-temperature Ising model (Lubetzky-Sly)
  - Random walk on Ramanujan graphs (Lubetzky-Peres)



Main Result

Definition

A p-shuffle: remove the top $Bin(N, p)$ cards from the deck and riffle as before.

Main Result

Definition

A p-shuffle: remove the top $Bin(N,p)$ cards from the deck and riffle as before.

Theorem

For $p \in (0,1)$, the p-shuffle mixes in $t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log N$ steps.

Main Result

Definition

A p-shuffle: remove the top $Bin(N, p)$ cards from the deck and riffle as before.

Theorem

For $p \in (0,1)$, the p-shuffle mixes in $t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log N$ steps.

- Definition of $\overline{C}_p$:

Definition

A p-shuffle: remove the top $Bin(N, p)$ cards from the deck and riffle as before.

Theorem

For $p \in (0,1)$, the p-shuffle mixes in $t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log N$ steps.

- Definition of $\overline{C}_p$:
  - With $q = 1 - p$, let $\theta_p \in [3,4)$ satisfy

$$p^{\theta_p} + q^{\theta_p} = (p^2 + q^2)^2.$$

Main Result

Definition

A p-shuffle: remove the top $Bin(N,p)$ cards from the deck and riffle as before.

Theorem

For $p \in (0,1)$, the p-shuffle mixes in $t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log N$ steps.

- Definition of $\overline{C}_p$:
    - With $q = 1 - p$, let $\theta_p \in [3,4)$ satisfy
$$p^{\theta_p} + q^{\theta_p} = (p^2 + q^2)^2.$$
    - Then set:
$$C_p = \frac{3 + \theta_p}{4 \log(1/(p^2 + q^2))},$$
$$\widetilde{C}_p = \frac{1}{\log(1/\max(p,q))},$$
$$\overline{C}_p = \max(C_p, \widetilde{C}_p)$$

Definition

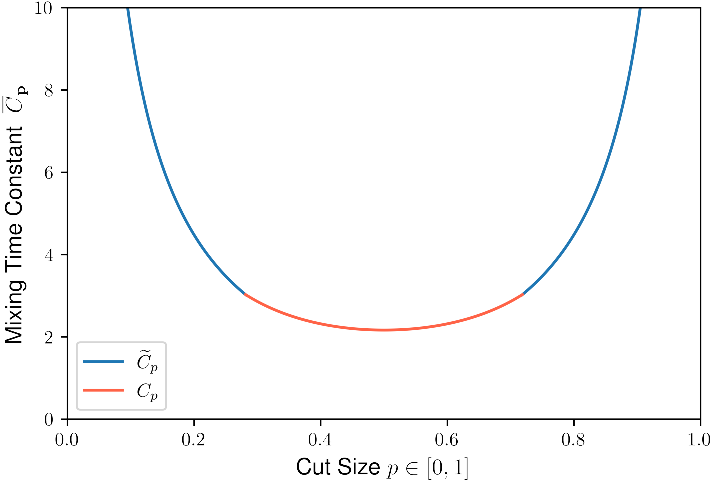
A p-shuffle: remove the top $Bin(N, p)$ cards from the deck and riffle as before.

Theorem

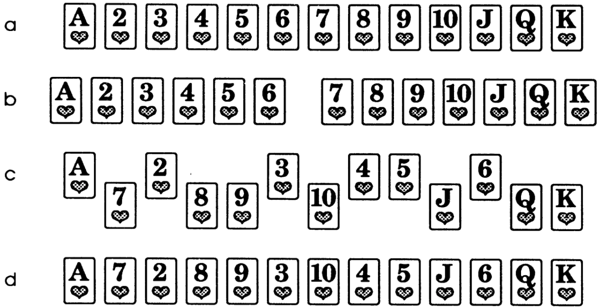
For $p \in (0,1)$, the p-shuffle mixes in $t_{\text{mix}} = (\overline{C}_p \pm o(1)) \log N$ steps.

- Definition of $\overline{C}_p$:
    - With $q = 1 - p$, let $\theta_p \in [3,4)$ satisfy
    $$p^{\theta_p} + q^{\theta_p} = (p^2 + q^2)^2.$$
    - Then set:
    $$C_p = \frac{3 + \theta_p}{4 \log(1/(p^2 + q^2))},$$
    $$\widetilde{C}_p = \frac{1}{\log(1/\max(p,q))},$$
    $$\overline{C}_p = \max(C_p, \widetilde{C}_p)$$
- Similar result for k-partite shuffles given any $(p_1, p_2, \cdots, p_k)$.

A Graph of $\overline{C}_p$

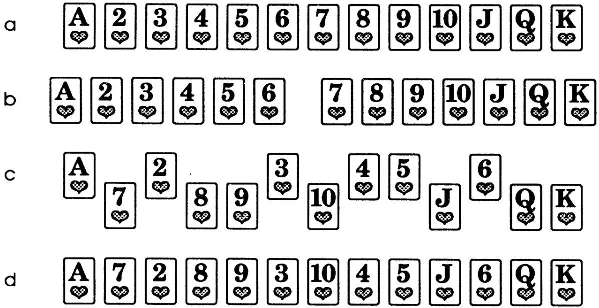


- After one riffle shuffle, the deck contains only 2 **rising sequences**.



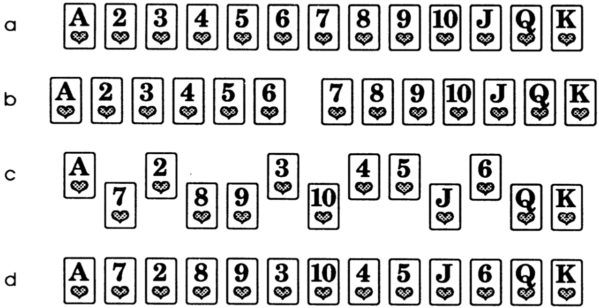
The Bayer-Diaconis Analysis

- After one riffle shuffle, the deck contains only 2 **rising sequences**.



- After t shuffles, $\leq 2^t$ rising sequences.

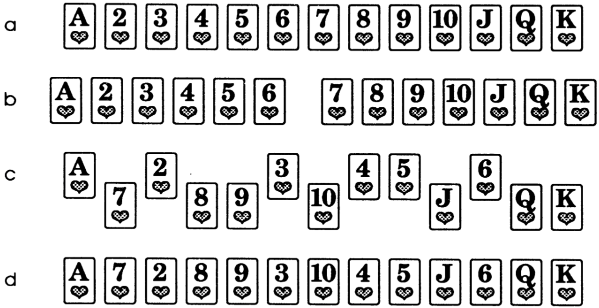
- After one riffle shuffle, the deck contains only 2 **rising sequences**.



- After t shuffles, $\leq 2^t$ rising sequences.
- A miracle: after t GSR shuffles, the deck distribution is **uniformly random** conditioned on the number of rising sequences.

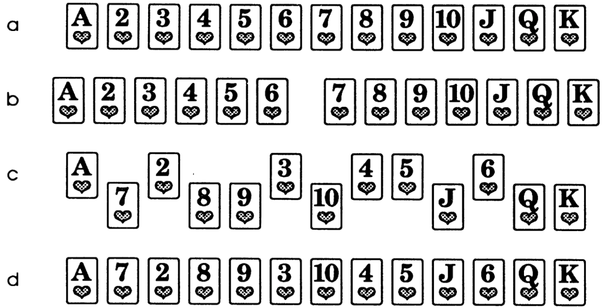
The Bayer-Diaconis Analysis

- After one riffle shuffle, the deck contains only 2 **rising sequences**.



- After t shuffles, $\leq 2^t$ rising sequences.
- A miracle: after t GSR shuffles, the deck distribution is **uniformly random** conditioned on the number of rising sequences.
  - Remains to analyze a 1 dimensional sufficient statistic.

- After one riffle shuffle, the deck contains only 2 **rising sequences**.



- After t shuffles, $\leq 2^t$ rising sequences.
- A miracle: after t GSR shuffles, the deck distribution is **uniformly random** conditioned on the number of rising sequences.
  - Remains to analyze a 1 dimensional sufficient statistic.
- With asymmetry, this miracle breaks. A new proof is needed.

Other Shuffles

- Many shuffling models have been studied:

Other Shuffles

- Many shuffling models have been studied:
  - Top-to-random: cutoff at $t_{\mathrm{mix}} = N \log N$ [Diaconis-Fill-Pitman 1992].

Other Shuffles

- Many shuffling models have been studied:
  - Top-to-random: cutoff at $t_{\mathrm{mix}} = N \log N$ [Diaconis-Fill-Pitman 1992].
  - Random-to-random: cutoff at $\frac{3}{4} N \log N$ [Subag 2013, Bernstein-Nestoridi 18].

Other Shuffles

- Many shuffling models have been studied:
  - Top-to-random: cutoff at $t_{\mathrm{mix}} = N \log N$ [Diaconis-Fill-Pitman 1992].
  - Random-to-random: cutoff at $\frac{3}{4} N \log N$ [Subag 2013, Bernstein-Nestoridi 18].
  - Cyclic-to-random: $t_{\mathrm{mix}} = \Theta(N \log N)$ [Pinsky 2015, Morris-Ning-Peres 2014].

- Many shuffling models have been studied:
  - Top-to-random: cutoff at $t_{\rm mix} = N \log N$ [Diaconis-Fill-Pitman 1992].
  - Random-to-random: cutoff at $\frac{3}{4} N \log N$ [Subag 2013, Bernstein-Nestoridi 18].
  - Cyclic-to-random: $t_{\rm mix} = \Theta(N \log N)$ [Pinsky 2015, Morris-Ning-Peres 2014].
  - Adjacent transpositions: cutoff at $t_{\rm mix} = \frac{N^3 \log N}{\pi^2}$ [Wilson 2004, Lacoin 2016].

- Many shuffling models have been studied:
  - Top-to-random: cutoff at $t_{\mathrm{mix}} = N \log N$ [Diaconis-Fill-Pitman 1992].
  - Random-to-random: cutoff at $\frac{3}{4} N \log N$ [Subag 2013, Bernstein-Nestoridi 18].
  - Cyclic-to-random: $t_{\mathrm{mix}} = \Theta(N \log N)$ [Pinsky 2015, Morris-Ning-Peres 2014].
  - Adjacent transpositions: cutoff at $t_{\mathrm{mix}} = \frac{N^3 \log N}{\pi^2}$ [Wilson 2004, Lacoin 2016].
  - Thorpe (perfect) riffle shuffle: $t_{\mathrm{mix}} \leq O(\log^3 N)$ [Montenegro-Tetali 2006, Morris 2005 & 2006 & 2008].

- Many shuffling models have been studied:
  - Top-to-random: cutoff at $t_{\mathrm{mix}} = N \log N$ [Diaconis-Fill-Pitman 1992].
  - Random-to-random: cutoff at $\frac{3}{4} N \log N$ [Subag 2013, Bernstein-Nestoridi 18].
  - Cyclic-to-random: $t_{\mathrm{mix}} = \Theta(N \log N)$ [Pinsky 2015, Morris-Ning-Peres 2014].
  - Adjacent transpositions: cutoff at $t_{\mathrm{mix}} = \frac{N^3 \log N}{\pi^2}$ [Wilson 2004, Lacoin 2016].
  - Thorpe (perfect) riffle shuffle: $t_{\mathrm{mix}} \leq O(\log^3 N)$ [Montenegro-Tetali 2006, Morris 2005 & 2006 & 2008].
  - Markovian riffling: $t_{\mathrm{mix}} \leq O(\log^4 N)$ [Jonasson-Morris 2015].

- [Bidigare-Hanlon-Rockmore 99, Brown-Diaconis 98, Stanley 01]: Eigenvalues are real, given by power sum symmetric functions. Connections to hyperplane arrangements.

- [Bidigare-Hanlon-Rockmore 99, Brown-Diaconis 98, Stanley 01]: Eigenvalues are real, given by power sum symmetric functions. Connections to hyperplane arrangements.

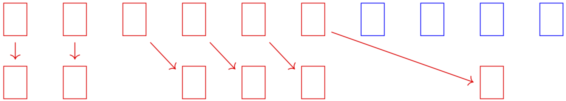
- [Assaf-Diaconis-Soundarajan 2011]: Cutoff with $O(1)$ window in L^∞ and separation distance. These are stricter notions of mixing involving worst case values of $\frac{P(x)}{Q(x)}$.

- [Bidigare-Hanlon-Rockmore 99, Brown-Diaconis 98, Stanley 01]: Eigenvalues are real, given by power sum symmetric functions. Connections to hyperplane arrangements.

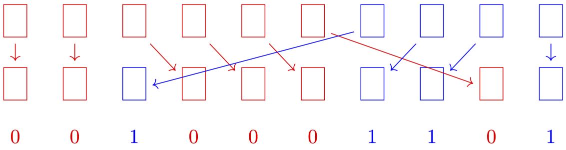
- [Assaf-Diaconis-Soundarajan 2011]: Cutoff with $O(1)$ window in L^∞ and separation distance. These are stricter notions of mixing involving worst case values of $\frac{P(x)}{Q(x)}$.

- [Lalley 2000]: Sharp lower bound for p close to $\{0, \frac{1}{2}, 1\}$. Identified the key **cold spots** phenomenon.

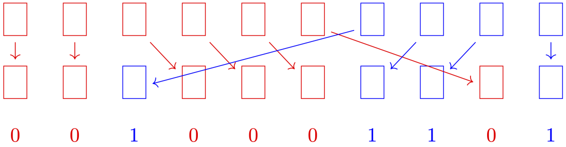
- Consider a single riffle shuffle:



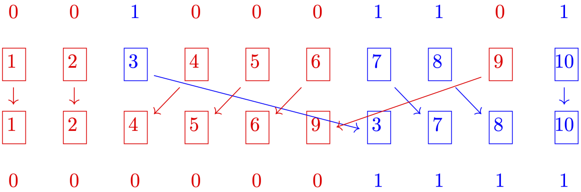
- Consider a single riffle shuffle:



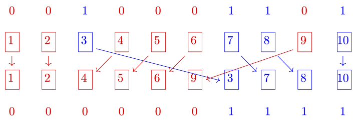
- Consider a single riffle shuffle:



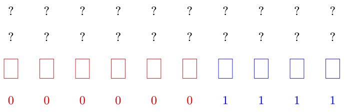
- To separate the two rising sequences, consider the inverse permutation (now with card labels).



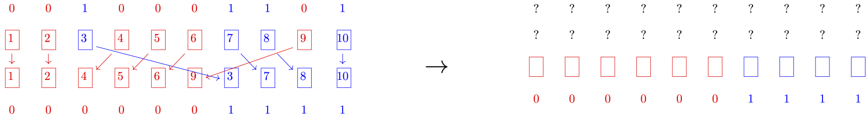
- Now, forget everything except the pile sizes.



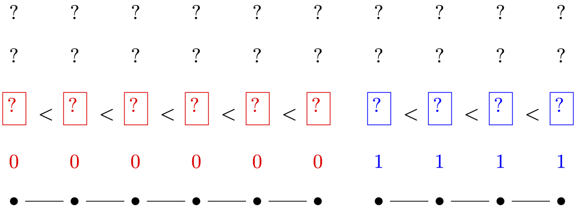
$\rightarrow$



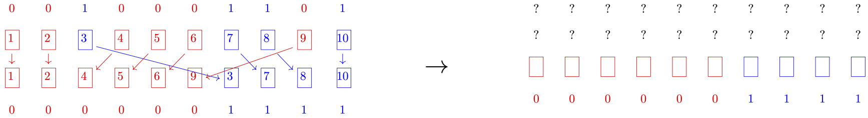
- Now, forget everything except the pile sizes.



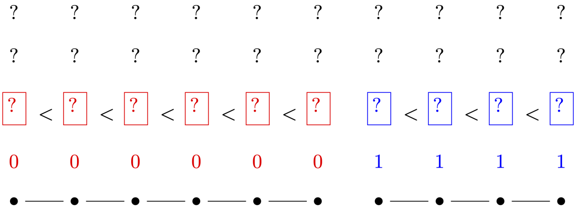
- All we remember is the split into increasing sequences:



- Now, forget everything except the pile sizes.



- All we remember is the split into increasing sequences:



- The conditional law is uniform given the constraints.

Transforming the Problem

- For t shuffles, 2^t card trajectories in $\{0,1\}^t$. Still uniform conditioned on the split:

$$\boxed{?} < \boxed{?} < \boxed{?} \qquad \boxed{?} < \boxed{?} < \boxed{?} \qquad \boxed{?} < \boxed{?} \qquad \boxed{?} < \boxed{?}$$

$$S = \quad 00 \quad 00 \quad 00 \quad 01 \quad 01 \quad 01 \quad 10 \quad 10 \quad 11 \quad 11$$

$$G = \quad \bullet \text{---} \bullet \text{---} \bullet \qquad \bullet \text{---} \bullet \text{---} \bullet \qquad \bullet \text{---} \bullet \qquad \bullet \text{---} \bullet$$

- For t shuffles, 2^t card trajectories in $\{0,1\}^t$. Still uniform conditioned on the split:

$$\boxed{?} < \boxed{?} < \boxed{?} \qquad \boxed{?} < \boxed{?} < \boxed{?} \qquad \boxed{?} < \boxed{?} \qquad \boxed{?} < \boxed{?}$$

$$S = 00 \quad 00 \quad 00 \quad 01 \quad 01 \quad 01 \quad 10 \quad 10 \quad 11 \quad 11$$

$$G = \bullet \text{---} \bullet \text{---} \bullet \qquad \bullet \text{---} \bullet \text{---} \bullet \qquad \bullet \text{---} \bullet \qquad \bullet \text{---} \bullet$$

- Conclusion: can generate the **inverse** of a p^{*t} shuffle as π^G below:

Transforming the Problem

- For t shuffles, 2^t card trajectories in $\{0,1\}^t$. Still uniform conditioned on the split:

$$\begin{array}{ccccccccccccc} & & \boxed{?} < & \boxed{?} < & \boxed{?} & \boxed{?} < & \boxed{?} < & \boxed{?} & \boxed{?} < & \boxed{?} & \boxed{?} < & \boxed{?} \\ S & = & 00 & 00 & 00 & 01 & 01 & 01 & 10 & 10 & 11 & 11 \\ G & = & \bullet - & \bullet - & \bullet & \bullet - & \bullet - & \bullet & \bullet - & \bullet & \bullet - & \bullet \end{array}$$

- Conclusion: can generate the **inverse** of a p^{*t} shuffle as π^G below:
  1. Generate N i.i.d. p-biased strings in $\{0,1\}^t$. Sort into increasing order:

$$S = (s_1, s_2, \cdots, s_N), \quad s_1 \le s_2 \le \cdots \le s_N.$$

$$S = (000, \quad 010, \quad 010, \quad 011, \quad 101, \quad 101, \quad 101, \quad 110, \quad 110, \quad 111)$$

- For t shuffles, 2^t card trajectories in $\{0,1\}^t$. Still uniform conditioned on the split:

$$\begin{array}{ccccccccccccc} & & \boxed{?} < \boxed{?} < \boxed{?} & & \boxed{?} < \boxed{?} < \boxed{?} & & \boxed{?} < \boxed{?} & & \boxed{?} < \boxed{?} \\ S & = & \color{red}{00} \quad \color{red}{00} \quad \color{red}{00} & & \color{red}{0}\color{blue}{1} \quad \color{red}{0}\color{blue}{1} \quad \color{red}{0}\color{blue}{1} & & \color{blue}{1}\color{red}{0} \quad \color{blue}{1}\color{red}{0} & & \color{blue}{11} \quad \color{blue}{11} \\ G & = & \bullet - \bullet - \bullet & & \bullet - \bullet - \bullet & & \bullet - \bullet & & \bullet - \bullet \end{array}$$

- Conclusion: can generate the **inverse** of a p^{*t} shuffle as π^G below:
  1. Generate N i.i.d. p-biased strings in $\{0,1\}^t$. Sort into increasing order:
  $$S = (s_1, s_2, \cdots, s_N), \quad s_1 \leq s_2 \leq \cdots \leq s_N.$$
  2. Connect $(i, i+1)$ if $s_i = s_{i+1}$, forming a graph $G = G(S)$.

$$\begin{array}{cccccccccccc} S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\ G & = & \bullet & \bullet & \!\!-\!\! \bullet & \bullet & \bullet & \!\!-\!\! \bullet & \!\!-\!\! \bullet & \bullet & \!\!-\!\! \bullet & \bullet \end{array}$$

- For t shuffles, 2^t card trajectories in $\{0,1\}^t$. Still uniform conditioned on the split:

$$\begin{array}{cccccccccccc}
 & & \boxed{?} & < \boxed{?} & < \boxed{?} & \boxed{?} & < \boxed{?} & < \boxed{?} & \boxed{?} & < \boxed{?} & \boxed{?} & < \boxed{?} \\
S & = & 00 & 00 & 00 & 01 & 01 & 01 & 10 & 10 & 11 & 11 \\
G & = & \bullet & \!\!-\!\!-\!\!-\, \bullet & \!\!-\!\!-\!\!-\, \bullet & \bullet & \!\!-\!\!-\!\!-\, \bullet & \!\!-\!\!-\!\!-\, \bullet & \bullet & \!\!-\!\!-\!\!-\, \bullet & \bullet & \!\!-\!\!-\!\!-\, \bullet
\end{array}$$

- Conclusion: can generate the **inverse** of a p^{*t} shuffle as π^G below:
  1. Generate N i.i.d. p-biased strings in $\{0,1\}^t$. Sort into increasing order:
  $$S = (s_1, s_2, \cdots, s_N), \quad s_1 \le s_2 \le \cdots \le s_N.$$
  2. Connect $(i, i+1)$ if $s_i = s_{i+1}$, forming a graph $G = G(S)$.
  3. Choose $\pi \in S_N$ uniformly. Sort π within G-components, forming $\pi^G \in S_N$.

$$\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet & \!\!-\!\!-\!\!-\, \bullet & \bullet & \bullet & \!\!-\!\!-\!\!-\, \bullet & \!\!-\!\!-\!\!-\, \bullet & \bullet & \!\!-\!\!-\!\!-\, \bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
 & & & & & & & & & \downarrow & \downarrow & \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}$$

Transforming the Problem

- For t shuffles, 2^t card trajectories in $\{0,1\}^t$. Still uniform conditioned on the split:

$$
\begin{array}{ccccccccccccc}
 & & \boxed{?} & < \boxed{?} & < \boxed{?} & \boxed{?} & < \boxed{?} & < \boxed{?} & \boxed{?} & < \boxed{?} & \boxed{?} & < \boxed{?} \\
S & = & 00 & 00 & 00 & 01 & 01 & 01 & 10 & 10 & 11 & 11 \\
G & = & \bullet & \!\!-\!\!\bullet & \!\!-\!\!\bullet & \bullet & \!\!-\!\!\bullet & \!\!-\!\!\bullet & \bullet & \!\!-\!\!\bullet & \bullet & \!\!-\!\!\bullet
\end{array}
$$

- Conclusion: can generate the **inverse** of a p^{*t} shuffle as π^G below:
  1. Generate N i.i.d. p-biased strings in $\{0,1\}^t$. Sort into increasing order:
     $$S = (s_1, s_2, \cdots, s_N), \quad s_1 \le s_2 \le \cdots \le s_N.$$
  2. Connect $(i, i+1)$ if $s_i = s_{i+1}$, forming a graph $G = G(S)$.
  3. Choose $\pi \in S_N$ uniformly. Sort π within G-components, forming $\pi^G \in S_N$.

$$
\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet & \!\!-\!\!\bullet & \bullet & \bullet & \!\!-\!\!\bullet & \!\!-\!\!\bullet & \bullet & \!\!-\!\!\bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
 & & & & & & & & & \downarrow & \downarrow & \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}
$$

- Inversion does not affect distance to uniformity. **How uniform is π^G?**

Easy Lower and Upper Bounds

- Mixed after t shuffles $\iff$ π^G is close to uniform.

$$
\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet & \!\!\text{---}\!\! \bullet & \bullet & \bullet & \!\!\text{---}\!\! \bullet & \!\!\text{---}\!\! \bullet & \bullet & \!\!\text{---}\!\! \bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
 & & & & & & & & & \downarrow & \downarrow & \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}
$$

- Mixed after t shuffles $\iff$ π^G is close to uniform.

$$
\begin{array}{ccccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet\;\text{---} & \bullet & \bullet & \bullet\;\text{---} & \bullet\;\text{---} & \bullet & \bullet\;\text{---} & \bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
 & & & & & & & & & \downarrow & \downarrow & \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}
$$

- Lower bound (tight for $p \geq 0.72$):

$$t_{\text{mix}} \geq \frac{\log N}{\log(1/p)}.$$

Easy Lower and Upper Bounds

- Mixed after t shuffles $\iff$ π^G is close to uniform.

$$
\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet & \!\!-\!\!-\!\!-\!\! \bullet & \bullet & \bullet & \!\!-\!\!-\!\!-\!\! \bullet & \!\!-\!\!-\!\!-\!\! \bullet & \bullet & \!\!-\!\!-\!\!-\!\! \bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
 & & & & & & & & & \downarrow & \downarrow & \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}
$$

- Lower bound (tight for $p \geq 0.72$):

$$t_{\text{mix}} \geq \frac{\log N}{\log(1/p)}.$$

  - If $t \leq \frac{(1-\varepsilon)\log N}{\log(1/p)}$, smallest N^δ strings are typically all zero:

$$s_1 = s_2 = \cdots = s_{N^\delta} = 0^t.$$

  - Then $\pi^G(1) < \pi^G(2) < \cdots < \pi^G(N^\delta)$. **Not a uniform permutation**.

Easy Lower and Upper Bounds

- Mixed after t shuffles $\iff$ π^G is close to uniform.

$$
\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet & \!\!\!\text{---}\!\!\! \bullet & \bullet & \bullet & \!\!\!\text{---}\!\!\! \bullet & \!\!\!\text{---}\!\!\! \bullet & \bullet & \!\!\!\text{---}\!\!\! \bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
 & & & & & & & & & \downarrow & \downarrow & \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}
$$

- Lower bound (tight for $p \geq 0.72$):

$$t_{\mathrm{mix}} \geq \frac{\log N}{\log(1/p)}.$$

  - If $t \leq \frac{(1-\varepsilon)\log N}{\log(1/p)}$, smallest N^δ strings are typically all zero:

$$s_1 = s_2 = \cdots = s_{N^\delta} = 0^t.$$

  - Then $\pi^G(1) < \pi^G(2) < \cdots < \pi^G(N^\delta)$. **Not a uniform permutation**.

- Upper bound (never tight):

$$t_{\mathrm{mix}} \leq \frac{2\log N}{\log\left(1/(p^2+q^2)\right)}.$$

Easy Lower and Upper Bounds

- Mixed after t shuffles $\iff$ π^G is close to uniform.

$$
\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 110, & 111) \\
G & = & \bullet & \bullet - & - \bullet & \bullet & \bullet - & - \bullet - & - \bullet & \bullet - & - \bullet & \bullet \\
\pi & = & 3 & 4 & 2 & 8 & 7 & 1 & 5 & 6 & 10 & 9 \\
\pi^G & = & 3 & 2 & 4 & 8 & 1 & 5 & 7 & 6 & 10 & 9
\end{array}
$$

- Lower bound (tight for $p \geq 0.72$):

$$t_{\mathrm{mix}} \geq \frac{\log N}{\log(1/p)}.$$

  - If $t \leq \frac{(1-\varepsilon)\log N}{\log(1/p)}$, smallest N^δ strings are typically all zero:

$$s_1 = s_2 = \cdots = s_{N^\delta} = 0^t.$$

  - Then $\pi^G(1) < \pi^G(2) < \cdots < \pi^G(N^\delta)$. **Not a uniform permutation**.

- Upper bound (never tight):

$$t_{\mathrm{mix}} \leq \frac{2\log N}{\log\left(1/(p^2+q^2)\right)}.$$

  - If $t \geq \frac{(2+\varepsilon)\log N}{\log(1/(p^2+q^2))}$, the strings $(s_1, \cdots, s_N)$ are typically **all distinct**.
  - On this event, G has no edges and $\pi^G = \pi$ is **uniform**.

Cold Spots: Another Obstruction to Mixing

- Lalley identified the constant $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$ based on **cold spots**.

Cold Spots: Another Obstruction to Mixing

- Lalley identified the constant $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$ based on **cold spots**.
- Suppose $x \in \{0,1\}^{\ell(x)}$ is a prefix with probability $\gg N^{-1/2}$.

Cold Spots: Another Obstruction to Mixing

- Lalley identified the constant $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$ based on **cold spots**.

- Suppose $x \in \{0,1\}^{\ell(x)}$ is a prefix with probability $\gg N^{-1/2}$.
- Strings $s_i, \dots, s_j$ with prefix x form an interval $I(x) = \{i, \cdots, j\} \subseteq [N]$ of length $|I(x)| \gg N^{1/2}$. The location of $I(x)$ is **essentially deterministic**.

Cold Spots: Another Obstruction to Mixing

- Lalley identified the constant $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$ based on **cold spots**.

- Suppose $x \in \{0,1\}^{\ell(x)}$ is a prefix with probability $\gg N^{-1/2}$.
- Strings $s_i, \ldots, s_j$ with prefix x form an interval $I(x) = \{i, \cdots, j\} \subseteq [N]$ of length $|I(x)| \gg N^{1/2}$. The location of $I(x)$ is **essentially deterministic**.

- With asymmetry, $\ell(x)$ can be larger if x's digits are very skewed.

- Lalley identified the constant $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$ based on **cold spots**.

- Suppose $x \in \{0,1\}^{\ell(x)}$ is a prefix with probability $\gg N^{-1/2}$.
- Strings $s_i, \ldots, s_j$ with prefix x form an interval $I(x) = \{i, \cdots, j\} \subseteq [N]$ of length $|I(x)| \gg N^{1/2}$. The location of $I(x)$ is **essentially deterministic**.

- With asymmetry, $\ell(x)$ can be larger if x's digits are very skewed.
- For $\ell(x)$ large, the local edge density of G within $I(x)$ is also large:

$$\mathbb{P}[(i, i+1) \in E(G)] \propto (p^2 + q^2)^{t-\ell(x)}.$$

Then $I(x)$ is a **cold spot**. Many G-edges $\implies \pi^G$ has extra ascents.

Cold Spots: Another Obstruction to Mixing

- Lalley identified the constant $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$ based on **cold spots**.

- Suppose $x \in \{0,1\}^{\ell(x)}$ is a prefix with probability $\gg N^{-1/2}$.
- Strings $s_i, \ldots, s_j$ with prefix x form an interval $I(x) = \{i, \cdots, j\} \subseteq [N]$ of length $|I(x)| \gg N^{1/2}$. The location of $I(x)$ is **essentially deterministic**.

- With asymmetry, $\ell(x)$ can be larger if x's digits are very skewed.
- For $\ell(x)$ large, the local edge density of G within $I(x)$ is also large:

$$\mathbb{P}[(i, i+1) \in E(G)] \propto (p^2 + q^2)^{t-\ell(x)}.$$

  Then $I(x)$ is a **cold spot**. Many G-edges $\implies \pi^G$ has extra ascents.

- Leads to a statistical test for π vs π^G:
  1. Fix a **digit profile** (c_0, c_1): prefix x must contain $c_0 \log N$ digits 0 and $c_1 \log N$ digits 1.
  2. Count **ascents** in the **cold spots** for all such x.
  3. Check if the number of ascents is typical for a uniform permutation.

- For $\sigma \in S_N$, define the ascent set

$$A(\sigma) = \{i \in [N-1] : \sigma(i+1) > \sigma(i)\}.$$

A Precise Heuristic from Ascents

- For $\sigma \in S_N$, define the ascent set

$$A(\sigma) = \{i \in [N-1] : \sigma(i+1) > \sigma(i)\}.$$

- Set $a_i = \mathbb{P}[(i, i+1) \in E(G)]$. With uniform $\pi \in S_N$, expect roughly

$$\mathbb{P}[i \in A(\pi)] = 1/2, \quad \mathbb{P}[i \in A(\pi^G)] = \frac{1 + a_i}{2}.$$

A Precise Heuristic from Ascents

- For $\sigma \in S_N$, define the ascent set

$$A(\sigma) = \{i \in [N-1] : \sigma(i+1) > \sigma(i)\}.$$

- Set $a_i = \mathbb{P}[(i, i+1) \in E(G)]$. With uniform $\pi \in S_N$, expect roughly

$$\mathbb{P}[i \in A(\pi)] = 1/2, \quad \mathbb{P}[i \in A(\pi^G)] = \frac{1+a_i}{2}.$$

- Let's pretend these events are **independent** over i.

A Precise Heuristic from Ascents

- For $\sigma \in S_N$, define the ascent set

$$A(\sigma) = \{i \in [N-1] : \sigma(i+1) > \sigma(i)\}.$$

- Set $a_i = \mathbb{P}[(i, i+1) \in E(G)]$. With uniform $\pi \in S_N$, expect roughly

$$\mathbb{P}[i \in A(\pi)] = 1/2, \quad \mathbb{P}[i \in A(\pi^G)] = \frac{1+a_i}{2}.$$

- Let's pretend these events are **independent** over i.
- Then for uniform $\sigma \in S_N$, likelihood ratio is random product

$$\frac{\mathbb{P}[\pi^G = \sigma]}{\mathbb{P}[\pi = \sigma]} = \prod_{i=1}^{N-1}(1 \pm a_i)$$

Behavior of an Independent Random Product

- Claim: if $\sum_i a_i^2 \ll 1$, then with high probability

$$\prod_{i=1}^{N-1} (1 \pm a_i) \approx 1.$$

Behavior of an Independent Random Product

- Claim: if $\sum_i a_i^2 \ll 1$, then with high probability

$$\prod_{i=1}^{N-1}(1 \pm a_i) \approx 1.$$

  - While ≈ 1, partial products form a martingale with QV $\approx a_i^2$.

Behavior of an Independent Random Product

- Claim: if $\sum_i a_i^2 \ll 1$, then with high probability

$$\prod_{i=1}^{N-1} (1 \pm a_i) \approx 1.$$

  - While ≈ 1, partial products form a martingale with QV $\approx a_i^2$.
- Claim: if $\sum_i a_i^2 \gg 1$, then with high probability

$$\prod_{i=1}^{N-1} (1 \pm a_i) \ll 1.$$

Behavior of an Independent Random Product

- Claim: if $\sum_i a_i^2 \ll 1$, then with high probability

$$\prod_{i=1}^{N-1}(1 \pm a_i) \approx 1.$$

  - While ≈ 1, partial products form a martingale with QV $\approx a_i^2$.

- Claim: if $\sum_i a_i^2 \gg 1$, then with high probability

$$\prod_{i=1}^{N-1}(1 \pm a_i) \ll 1.$$

  - Follows from LLN and $\mathbb{E}[\log(1 \pm a_i)] \leq 1 - \frac{a_i^2}{2}$.

- Claim: if $\sum_i a_i^2 \ll 1$, then with high probability

$$\prod_{i=1}^{N-1}(1 \pm a_i) \approx 1.$$

  - While ≈ 1, partial products form a martingale with QV $\approx a_i^2$.
- Claim: if $\sum_i a_i^2 \gg 1$, then with high probability

$$\prod_{i=1}^{N-1}(1 \pm a_i) \ll 1.$$

  - Follows from LLN and $\mathbb{E}[\log(1 \pm a_i)] \leq 1 - \frac{a_i^2}{2}$.
- Convenient observation: with G' an independent copy of G,

$$\sum_i a_i^2 = \mathbb{E}\left[\left|E(G) \cap E(G')\right|\right].$$

Behavior of an Independent Random Product

- Claim: if $\sum_i a_i^2 \ll 1$, then with high probability

$$\prod_{i=1}^{N-1} (1 \pm a_i) \approx 1.$$

  - While ≈ 1, partial products form a martingale with QV $\approx a_i^2$.
- Claim: if $\sum_i a_i^2 \gg 1$, then with high probability

$$\prod_{i=1}^{N-1} (1 \pm a_i) \ll 1.$$

  - Follows from LLN and $\mathbb{E}[\log(1 \pm a_i)] \le 1 - \frac{a_i^2}{2}$.
- Convenient observation: with G' an independent copy of G,

$$\sum_i a_i^2 = \mathbb{E}\left[\left|E(G) \cap E(G')\right|\right].$$

- Write $E(G, G') = E(G) \cap E(G')$ for the set of shared edges. Expect:

$$\text{Mixed after } t \text{ shuffles} \iff \mathbb{E}[|E(G, G')|] \ll 1.$$

- When the first moment

$$\mathbb{E}[|E(G, G')|] \gg N^{\delta}$$

is large, **not mixed**.

- When the first moment

$$\mathbb{E}[|E(G,G')|] \gg N^{\delta}$$

  is large, **not mixed**.

- When (truncated) **exponential** moment

$$\mathbb{E}\left[e^{c\cdot|E(G,G')|}\right] \leq 1 + O(N^{-\delta})$$

  is small for bounded c, **mixed**.

The True Picture

- When the first moment

$$\mathbb{E}[|E(G,G')|] \gg N^{\delta}$$

  is large, **not mixed**.

- When (truncated) **exponential** moment

$$\mathbb{E}\left[e^{c\cdot|E(G,G')|}\right] \leq 1 + O(N^{-\delta})$$

  is small for bounded c, **mixed**.

- Fortunately, these criteria match: $|E(G,G')|$ transitions from $\gg 1$ to $\ll 1$ almost **simultaneously** in 1st and exponential moment senses.
  - (Not quite true for k-partite shuffles.)

Proof Overview

- Three main components in the proof:
  1. Show mixing if $|E(G, G')|$ has small exponential moments:

  $$\mathbb{E}[e^{c|E(G,G')|}] \leq 1 + N^{-\delta}.$$

  (Up to some truncation.)

Proof Overview

- Three main components in the proof:
  1. Show mixing if $|E(G,G')|$ has small exponential moments:
  $$\mathbb{E}[e^{c|E(G,G')|}] \le 1 + N^{-\delta}.$$
  (Up to some truncation.)
  2. Reduce **exponential** moment to **first** moment control
  $$\mathbb{E}[|E(G,G')|] \le N^{-\delta}.$$

Proof Overview

- Three main components in the proof:
  1. Show mixing if $|E(G,G')|$ has small exponential moments:
  $$\mathbb{E}[e^{c|E(G,G')|}] \le 1 + N^{-\delta}.$$
  (Up to some truncation.)
  2. Reduce **exponential** moment to **first** moment control
  $$\mathbb{E}[|E(G,G')|] \le N^{-\delta}.$$
  3. Understand first moment $\implies$ upper bound t_{mix}.

Proof Overview

- Three main components in the proof:
  1. Show mixing if $|E(G, G')|$ has small exponential moments:
  $$\mathbb{E}[e^{c|E(G,G')|}] \le 1 + N^{-\delta}.$$
  (Up to some truncation.)
  2. Reduce **exponential** moment to **first** moment control
  $$\mathbb{E}[|E(G, G')|] \le N^{-\delta}.$$
  3. Understand first moment $\implies$ upper bound t_{mix}.
     - Main contribution to first moment $\implies$ optimal choice of cold spots to lower bound t_{mix}.

χ^2 Upper Bound on TV Distance

- To upper-bound distance from uniformity, use the "χ^2 trick".

χ^2 Upper Bound on TV Distance

- To upper-bound distance from uniformity, use the "χ^2 trick".
- Let $\mathbf{F}(\sigma) = N! \cdot \mathbb{P}_{\pi \sim U(S_n), G \sim \mathcal{G}}[\pi^G = \sigma]$. TV distance to uniformity is

$$\mathbb{E}_{\sigma \sim U(S_n)}[|\mathbf{F}(\sigma) - 1|].$$

χ^2 Upper Bound on TV Distance

- To upper-bound distance from uniformity, use the "χ^2 trick".
- Let $\mathbf{F}(\sigma) = N! \cdot \mathbb{P}_{\pi\sim U(S_n), G\sim\mathcal{G}}[\pi^G = \sigma]$. TV distance to uniformity is

$$\mathbb{E}_{\sigma\sim U(S_n)}[|\mathbf{F}(\sigma) - 1|].$$

- As $\mathbb{E}_{\sigma\sim U(S_n)}[\mathbf{F}(\sigma)] = 1$, Cauchy-Schwarz gives

$$\left(\mathbb{E}_{\sigma\sim U(S_n)}|\mathbf{F}(\sigma) - 1|\right)^2 \leq \mathbb{E}_{\sigma\sim U(S_n)}[\mathbf{F}(\sigma)^2 - 1] \overset{?}{\ll} 1.$$

χ^2 Upper Bound on TV Distance

- To upper-bound distance from uniformity, use the "χ^2 trick".
- Let $\mathbf{F}(\sigma) = N! \cdot \mathbb{P}_{\pi \sim U(S_n), G \sim \mathcal{G}}[\pi^G = \sigma]$. TV distance to uniformity is

$$\mathbb{E}_{\sigma \sim U(S_n)}[|\mathbf{F}(\sigma) - 1|].$$

- As $\mathbb{E}_{\sigma \sim U(S_n)}[\mathbf{F}(\sigma)] = 1$, Cauchy-Schwarz gives

$$\left(\mathbb{E}_{\sigma \sim U(S_n)}|\mathbf{F}(\sigma) - 1|\right)^2 \leq \mathbb{E}_{\sigma \sim U(S_n)}[\mathbf{F}(\sigma)^2 - 1] \overset{?}{\ll} 1.$$

- Let (π', G') be an independent copy. Then

$$\mathbf{F}(\sigma)^2 = (N!)^2 \cdot \mathbb{P}[\pi^G = \sigma, (\pi')^{G'} = \sigma]$$

$$\implies \mathbb{E}_{\sigma \sim U(S_n)}[\mathbf{F}(\sigma)^2] = N! \cdot \mathbb{P}[\pi^G = (\pi')^{G'}].$$

χ^2 Upper Bound on TV Distance

- To upper-bound distance from uniformity, use the "χ^2 trick".
- Let $\mathbf{F}(\sigma) = N! \cdot \mathbb{P}_{\pi \sim U(S_n), G \sim \mathcal{G}}[\pi^G = \sigma]$. TV distance to uniformity is

$$\mathbb{E}_{\sigma \sim U(S_n)}[|\mathbf{F}(\sigma) - 1|].$$

- As $\mathbb{E}_{\sigma \sim U(S_n)}[\mathbf{F}(\sigma)] = 1$, Cauchy-Schwarz gives

$$\left(\mathbb{E}_{\sigma \sim U(S_n)}|\mathbf{F}(\sigma) - 1|\right)^2 \leq \mathbb{E}_{\sigma \sim U(S_n)}[\mathbf{F}(\sigma)^2 - 1] \overset{?}{\ll} 1.$$

- Let (π', G') be an independent copy. Then

$$\mathbf{F}(\sigma)^2 = (N!)^2 \cdot \mathbb{P}[\pi^G = \sigma, (\pi')^{G'} = \sigma]$$
$$\implies \mathbb{E}_{\sigma \sim U(S_n)}[\mathbf{F}(\sigma)^2] = N! \cdot \mathbb{P}[\pi^G = (\pi')^{G'}].$$

- Define

$$f_{G,G'} = N! \cdot \mathbb{P}_{\pi, \pi' \sim U(S_n)}[\pi^G = (\pi')^{G'} | G, G'].$$

$f_{G,G'}$ measures "interaction" between G and G'. We'll try to show:

$$\mathbb{E}_{G,G' \sim \mathcal{G}}[|f_{G,G'} - 1|] \overset{?}{\approx} 0.$$

- Suppose G, G' have no shared or adjacent edges.

$$
\begin{array}{cccccccccccc}
G & = & \bullet & \bullet & \!\!\!\text{---}\!\!\! \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \!\!\!\text{---}\!\!\! \bullet & \bullet \\
\pi & = & 9 & 10 & 4 & 7 & 6 & 3 & 8 & 2 & 1 & 5 \\
\pi^G & = & 9 & 4 & 10 & 7 & 6 & 3 & 8 & 1 & 2 & 5 \\
G' & = & \bullet & \bullet & \bullet & \bullet & \bullet & \!\!\!\text{---}\!\!\! \bullet & \!\!\!\text{---}\!\!\! \bullet & \bullet & \bullet & \bullet \\
\pi' & = & 10 & 3 & 1 & 9 & 2 & 8 & 6 & 5 & 7 & 4 \\
 & & & & & & \downarrow & & & & & \\
(\pi')^{G'} & = & 10 & 3 & 1 & 9 & 2 & 6 & 8 & 5 & 7 & 4
\end{array}
$$

Reducing Upper Bound to Exponential Moment

- Suppose G, G' have no shared or adjacent edges.

$$
\begin{array}{ccccccccccccccccccccc}
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet \\
\pi & = & 9 & & 10 & & 4 & & 7 & & 6 & & 3 & & 8 & & 2 & & 1 & & 5 \\
\pi^G & = & 9 & & 4 & & 10 & & 7 & & 6 & & 3 & & 8 & & 1 & & 2 & & 5 \\
G' & = & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet \\
\pi' & = & 10 & & 3 & & 1 & & 9 & & 2 & & 8 & & 6 & & 5 & & 7 & & 4 \\
(\pi')^{G'} & = & 10 & & 3 & & 1 & & 9 & & 2 & & 6 & & 8 & & 5 & & 7 & & 4
\end{array}
$$

- Then the transformations $(\cdot)^G$ and $(\cdot)^{G'}$ "don't interact". So,

$$f_{G,G'} = N! \cdot \mathbb{P}_{\pi,\pi'}[\pi^G = (\pi')^{G'} | G, G'] = 1$$

as if $\pi^G, (\pi')^{G'}$ were uniform and independent. Good so far...

- Now suppose G, G' share a single edge $(i, i+1)$.

$$
\begin{array}{cccccccccccc}
G & = & \bullet & \bullet\,\text{—} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet\,\text{—} & \bullet & \bullet \\
\pi & = & 9 & 10 & 4 & 7 & 6 & 3 & 8 & 2 & 1 & 5 \\
\pi^G & = & 9 & 4 & 10 & 7 & 6 & 3 & 8 & 1 & 2 & 5 \\
G' & = & \bullet & \bullet\,\text{—} & \bullet & \bullet & \bullet\,\text{—} & \bullet\,\text{—} & \bullet & \bullet & \bullet & \bullet \\
\pi' & = & 10 & 3 & 1 & 9 & 2 & 8 & 6 & 5 & 7 & 4 \\
(\pi')^{G'} & = & 10 & 1 & 3 & 9 & 2 & 6 & 8 & 5 & 7 & 4
\end{array}
$$

- Now suppose G, G' share a single edge $(i, i+1)$.

$$
\begin{array}{ccccccccccccccccccccc}
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet \\
\pi & = & 9 & & 10 & & 4 & & 7 & & 6 & & 3 & & 8 & & 2 & & 1 & & 5 \\
\pi^G & = & 9 & & 4 & & 10 & & 7 & & 6 & & 3 & & 8 & & 1 & & 2 & & 5 \\
G' & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet \\
\pi' & = & 10 & & 3 & & 1 & & 9 & & 2 & & 8 & & 6 & & 5 & & 7 & & 4 \\
(\pi')^{G'} & = & 10 & & 1 & & 3 & & 9 & & 2 & & 6 & & 8 & & 5 & & 7 & & 4
\end{array}
$$

- Then $(i, i+1)$ is always an ascent for π^G and $(\pi')^{G'}$. Result:

$$f_{G,G'} = N! \cdot \mathbb{P}_{\pi,\pi'}[\pi^G = (\pi')^{G'}] = \mathbf{2}.$$

Reducing Upper Bound to Exponential Moment

- Now suppose G, G' share a single edge $(i, i+1)$.

$$
\begin{array}{cccccccccccc}
G & = & \bullet & \bullet & \!\!\text{---}\!\! \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \!\!\text{---}\!\! \bullet & \bullet \\
\pi & = & 9 & 10 & 4 & 7 & 6 & 3 & 8 & 2 & 1 & 5 \\
\pi^G & = & 9 & 4 & 10 & 7 & 6 & 3 & 8 & 1 & 2 & 5 \\
G' & = & \bullet & \bullet & \!\!\text{---}\!\! \bullet & \bullet & \bullet & \!\!\text{---}\!\! \bullet & \!\!\text{---}\!\! \bullet & \bullet & \bullet & \bullet \\
\pi' & = & 10 & 3 & 1 & 9 & 2 & 8 & 6 & 5 & 7 & 4 \\
(\pi')^{G'} & = & 10 & 1 & 3 & 9 & 2 & 6 & 8 & 5 & 7 & 4
\end{array}
$$

- Then $(i, i+1)$ is always an ascent for π^G and $(\pi')^{G'}$. Result:

$$f_{G,G'} = N! \cdot \mathbb{P}_{\pi,\pi'}[\pi^G = (\pi')^{G'}] = \mathbf{2}.$$

- Disjoint interactions between G, G' combine **multiplicatively**.
- Assuming "constant diameter" interactions (via truncation):

$$f_{G,G'} \leq \mathbf{e}^{\mathbf{c}|\mathbf{E}(\mathbf{G},\mathbf{G}')|}.$$

1. Show mixing if $|E(G,G')|$ has small truncated exponential moments:

$$\mathbb{E}[e^{c|E(G,G')|}] \leq 1 + N^{-\delta}.$$

2. Reduce exponential moments to the first moment bound

$$\mathbb{E}[|E(G,G')|] \leq N^{-\delta}.$$

3. Understand first moment $\implies$ upper bound t_{mix}.
   - Main contribution to first moment $\implies$ optimal choice of cold spots to lower bound t_{mix}.

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j\geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

$$\begin{array}{cccccccccccc}
S & = & (000, & 010, & ???, & ???, & ???, & ???, & ???, & ???, & ???, & ???) \\
G & = & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
S' & = & (001, & 001, & ???, & ???, & ???, & ???, & ???, & ???, & ???, & ???) \\
G' & = & \bullet & \!\!\!\!\text{———}\!\!\!\! \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
E(G, G') & = & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet
\end{array}$$

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

$$\begin{array}{ccccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & ???, & & ???, & & ???, & & ???, & & ???, & & ???, & & ???) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\
S' & = & (001, & & 001, & & 001, & & ???, & & ???, & & ???, & & ???, & & ???, & & ???, & & ???) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\
E(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet
\end{array}$$

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G,G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G,G')$ is unlikely to have more edges.

$$\begin{array}{cccccccccccc}
S & = & (000, & 010, & 010, & 011, & ???, & ???, & ???, & ???, & ???, & ???) \\
G & = & \bullet & \bullet \!\!\text{———}\!\! & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
S' & = & (001, & 001, & 001, & 011, & ???, & ???, & ???, & ???, & ???, & ???) \\
G' & = & \bullet \!\!\text{———}\!\! & \bullet \!\!\text{———}\!\! & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
E(G,G') & = & \bullet & \bullet \!\!\text{———}\!\! & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet
\end{array}$$

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

$$\begin{array}{cccccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & 011, & & 101, & & ???, & & ???, & & ???, & & ???, & & ???) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\
S' & = & (001, & & 001, & & 001, & & 011, & & 011, & & ???, & & ???, & & ???, & & ???, & & ???) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\
E(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet & & \bullet
\end{array}$$

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

$$
\begin{array}{ccllllllllll}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & ???, & ???, & ???, & ???) \\
G & = & \bullet & \bullet \text{———} & \bullet & \bullet & \bullet \text{———} & \bullet & \bullet & \bullet & \bullet & \bullet \\
S' & = & (001, & 001, & 001, & 011, & 011, & 010, & ???, & ???, & ???, & ???) \\
G' & = & \bullet \text{———} & \bullet \text{———} & \bullet & \bullet \text{———} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
E(G,G') & = & \bullet & \bullet \text{———} & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet
\end{array}
$$

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

$$\begin{array}{ccccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & 011, & & 101, & & 101, & & 101, & & ???, & & ???, & & ???) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet \\
S' & = & (001, & & 001, & & 001, & & 011, & & 011, & & 010, & & 010, & & ???, & & ???, & & ???) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet \\
E(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet
\end{array}$$

Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

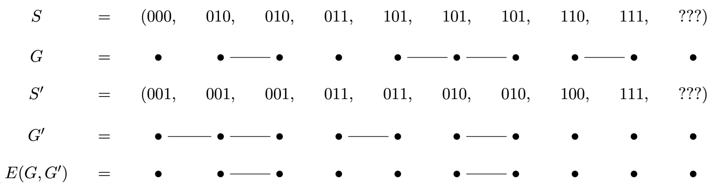
  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

$$\begin{array}{ccccccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & 011, & & 101, & & 101, & & 101, & & 110, & & ???, & & ???) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet \\
S' & = & (001, & & 001, & & 001, & & 011, & & 011, & & 010, & & 010, & & 100, & & ???, & & ???) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet \\
E(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet
\end{array}$$

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.

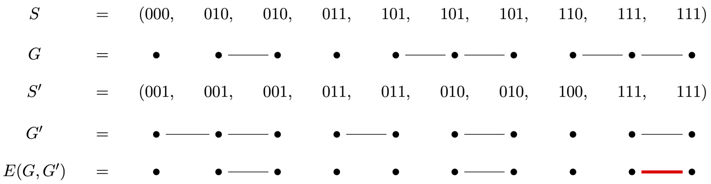


Hazard Rate Method to Bound Exponential Moments

- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

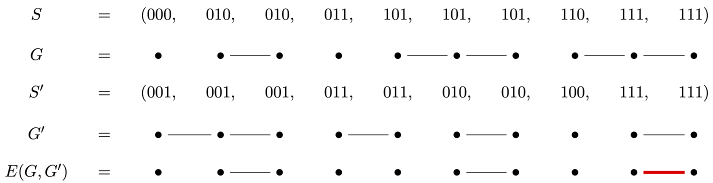
  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.



- Suppose random variable $X \in \mathbb{Z}_{\geq 0}$ has hazard rate uniformly close to 1:

$$\sup_{j \geq 0} \mathbb{P}[X \geq j+1 | X \geq j] \leq \varepsilon.$$

  - Then $\mathbb{E}[e^{cX}] = 1 + O(c\varepsilon)$ follows for $c \ll \varepsilon^{-1}$.
- Try taking $X = |E(G, G')|$. Explore $(s_1, s_1'), (s_2, s_2'), \cdots (s_N, s_N')$ in order.
- Hope: at **any time**, $E(G, G')$ is unlikely to have more edges.



- Uh oh! If $s_j = s_j' = 111 \cdots 1$, future edges are guaranteed to be in $E(G, G')$. The exploration "ran out of space".

- Fix: explore both forward and backward. **Stop exploration early**.

$$
\begin{array}{lcccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & 011, & & 101, & & 101, & & 101, & & 110, & & 111, & & 111) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet \\
S' & = & (001, & & 001, & & 001, & & 011, & & 011, & & 010, & & 010, & & 100, & & 111, & & 111) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet \\
E(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet
\end{array}
$$

- Fix: explore both forward and backward. **Stop exploration early**.

$$
\begin{array}{lcllllllllll}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 111, & 111) \\
G & = & \bullet & \bullet - & \bullet & \bullet & \bullet - & \bullet - & \bullet & \bullet - & \bullet - & \bullet \\
S' & = & (001, & 001, & 001, & 011, & 011, & 010, & 010, & 100, & 111, & 111) \\
G' & = & \bullet - & \bullet - & \bullet & \bullet - & \bullet & \bullet - & \bullet & \bullet & \bullet - & \bullet \\
E(G,G') & = & \bullet & \bullet - & \bullet & \bullet & \bullet & \bullet - & \bullet & \bullet & \bullet - & \bullet \\
E_{for}(G,G') & = & \bullet & \bullet - & \bullet & \bullet & \bullet & \bullet - & \bullet & & &
\end{array}
$$

- Fix: explore both forward and backward. **Stop exploration early**.

$$
\begin{array}{rcccccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & 011, & & 101, & & 101, & & 101, & & 110, & & 111, & & 111) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet \\
S' & = & (001, & & 001, & & 001, & & 011, & & 011, & & 010, & & 010, & & 100, & & 111, & & 111) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet \\
E(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet \\
E_{for}(G,G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & & & & & \\
E_{back}(G,G') & = & & & & & & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet
\end{array}
$$

Forward-Backward Covering of $E(G, G')$

- Fix: explore both forward and backward. **Stop exploration early.**

$$
\begin{array}{lcllllllllll}
S & = & (000, & 010, & 010, & 011, & 101, & 101, & 101, & 110, & 111, & 111) \\
G & = & \bullet & \bullet \text{ --- } & \bullet & \bullet & \bullet \text{ --- } & \bullet \text{ --- } & \bullet & \bullet \text{ --- } & \bullet \text{ --- } & \bullet \\
S' & = & (001, & 001, & 001, & 011, & 011, & 010, & 010, & 100, & 111, & 111) \\
G' & = & \bullet \text{ --- } & \bullet \text{ --- } & \bullet & \bullet \text{ --- } & \bullet & \bullet \text{ --- } & \bullet & \bullet & \bullet \text{ --- } & \bullet \\
E(G,G') & = & \bullet & \bullet \text{ --- } & \bullet & \bullet & \bullet & \bullet \text{ --- } & \bullet & \bullet & \bullet \text{ --- } & \bullet \\
E_{for}(G,G') & = & \bullet & \bullet \text{ --- } & \bullet & \bullet & \bullet & \bullet \text{ --- } & \bullet & & & \\
E_{back}(G,G') & = & & & & \bullet & \bullet & \bullet \text{ --- } & \bullet & \bullet & \bullet \text{ --- } & \bullet
\end{array}
$$

- Stop forward exploration when prefix **11** appears. Backward, stop on **00**.

- Fix: explore both forward and backward. **Stop exploration early**.

$$
\begin{aligned}
S &= (000,\ 010,\ 010,\ 011,\ 101,\ 101,\ 101,\ \mathbf{11}0,\ 111,\ 111) \\
G &= \bullet \quad \bullet - \bullet \quad \bullet \quad \bullet - \bullet - \bullet \quad \bullet - \bullet - \bullet \\
S' &= (001,\ 001,\ \mathbf{00}1,\ 011,\ 011,\ 010,\ 010,\ 100,\ 111,\ 111) \\
G' &= \bullet - \bullet - \bullet \quad \bullet - \bullet \quad \bullet - \bullet \quad \bullet \quad \bullet - \bullet \\
E(G,G') &= \bullet \quad \bullet - \bullet \quad \bullet \quad \bullet \quad \bullet - \bullet \quad \bullet \quad \bullet - \bullet \\
E_{for}(G,G') &= \bullet \quad \bullet - \bullet \quad \bullet \quad \bullet \quad \bullet - \bullet \\
E_{back}(G,G') &= \bullet \quad \bullet \quad \bullet - \bullet \quad \bullet \quad \bullet - \bullet
\end{aligned}
$$

- Stop forward exploration when prefix **11** appears. Backward, stop on **00**.
- Use hazard rate method on $|E_{for}(G, G')|$ and $|E_{back}(G, G')|$ separately.

Forward-Backward Covering of $E(G, G')$

- Fix: explore both forward and backward. **Stop exploration early**.

$$
\begin{array}{cccccccccccccccccccccc}
S & = & (000, & & 010, & & 010, & & 011, & & 101, & & 101, & & 101, & & \mathbf{11}0, & & 111, & & 111) \\
G & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & \text{---} & \bullet \\
S' & = & (001, & & 001, & & \mathbf{00}1, & & 011, & & 011, & & 010, & & 010, & & 100, & & 111, & & 111) \\
G' & = & \bullet & \text{---} & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet \\
E(G, G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet \\
E_{for}(G, G') & = & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & & & & & \\
E_{back}(G, G') & = & & & & & & & \bullet & & \bullet & & \bullet & \text{---} & \bullet & & \bullet & & \bullet & \text{---} & \bullet
\end{array}
$$

- Stop forward exploration when prefix **11** appears. Backward, stop on **00**.
- Use hazard rate method on $|E_{for}(G, G')|$ and $|E_{back}(G, G')|$ separately.
- Can ensure $E(G, G') = E_{for}(G, G') \cup E_{back}(G, G')$ by truncation. Then

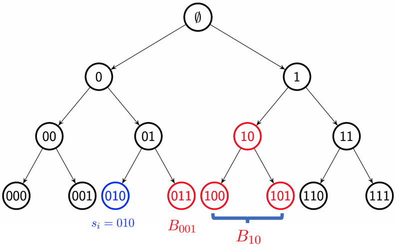
$$e^{c|E(G,G')|} \leq \left(e^{2c|E_{for}(G,G')|} + e^{2c|E_{back}(G,G')|}\right)/2 \approx 1.$$

Conditional Behavior of Forward Exploration

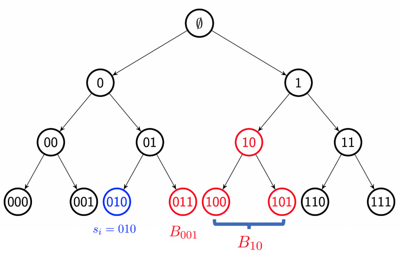
- Want to control unrevealed edges in $E_{for}(G, G')$ under conditioning.

- Want to control unrevealed edges in $E_{for}(G, G')$ under conditioning.
- Given s_i, relevant strings lie in "lexicographic subinterval" of $\{0, 1\}^t$ between s_i and **11**.

- Want to control unrevealed edges in $E_{for}(G, G')$ under conditioning.
- Given s_i, relevant strings lie in "lexicographic subinterval" of $\{0,1\}^t$ between s_i and **11**.
- Partition this subinterval into $O(\log N) \ll N^\delta$ **prefix blocks** B_x.

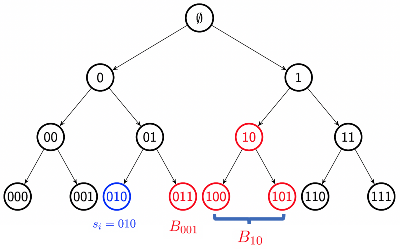


- Want to control unrevealed edges in $E_{for}(G, G')$ under conditioning.
- Given s_i, relevant strings lie in "lexicographic subinterval" of $\{0,1\}^t$ between s_i and **11**.
- Partition this subinterval into $O(\log N) \ll N^{\delta}$ **prefix blocks** B_x.



- The conditional problem reduces to smaller versions of the original problem within each block B_x, with $t - \ell(x)$ unassigned digits.

- Want to control unrevealed edges in $E_{for}(G, G')$ under conditioning.
- Given s_i, relevant strings lie in "lexicographic subinterval" of $\{0,1\}^t$ between s_i and **11**.
- Partition this subinterval into $O(\log N) \ll N^\delta$ **prefix blocks** B_x.



- The conditional problem reduces to smaller versions of the original problem within each block B_x, with $t - \ell(x)$ unassigned digits.
- By **early stopping**, the conditional law for the number of strings landing in some B_x can never blow up much.

1. Show mixing if $|E(G,G')|$ has small truncated exponential moments:

$$\mathbb{E}[e^{c|E(G,G')|}] \leq 1 + N^{-\delta}.$$

2. Reduce exponential moment estimate to first moment control

$$\mathbb{E}[|E(G,G')|] \leq N^{-\delta}.$$

3. Understand first moment $\implies$ upper bound t_{mix}.

1. Show mixing if $|E(G,G')|$ has small truncated exponential moments:

$$\mathbb{E}[e^{c|E(G,G')|}] \le 1 + N^{-\delta}.$$

2. Reduce exponential moment estimate to first moment control

$$\mathbb{E}[|E(G,G')|] \le N^{-\delta}.$$

3. Understand first moment $\implies$ upper bound t_{mix}.

- Main contribution to first moment $\implies$ optimal choice of cold spots to lower bound t_{mix}.

Controlling the First Moment

- How to understand $\mathbb{E}[|E(G,G')|]$? With $a_i = \mathbb{P}[(i,i+1) \in E(G)]$,

$$\mathbb{E}[|E(G,G')|] = \sum_{i=1}^{N-1} a_i^2.$$

Need to understand the values a_i.

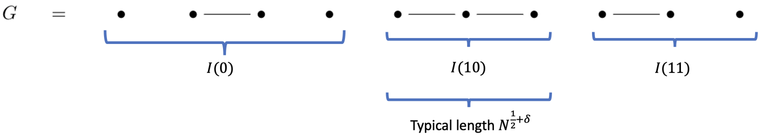
- How to understand $\mathbb{E}[|E(G, G')|]$? With $a_i = \mathbb{P}[(i, i+1) \in E(G)]$,

$$\mathbb{E}[|E(G, G')|] = \sum_{i=1}^{N-1} a_i^2.$$

Need to understand the values a_i.

- Partition $\{0, 1\}^t$ into certain prefix blocks $\{B_x : x \in \mathcal{L}\}$.
- Partition $[N]$ into discrete intervals $I(x)$ of strings with prefix x.

$$S = (000, \quad 010, \quad 010, \quad 011, \quad 101, \quad 101, \quad 101, \quad 110, \quad 110, \quad 111)$$

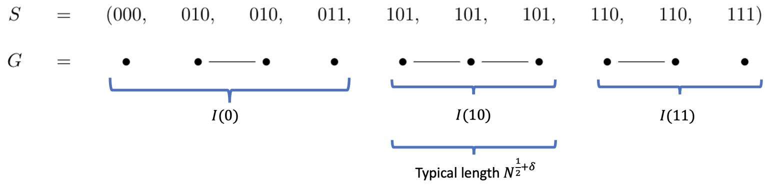


Controlling the First Moment

- How to understand $\mathbb{E}[|E(G,G')|]$? With $a_i = \mathbb{P}[(i,i+1) \in E(G)]$,

$$\mathbb{E}[|E(G,G')|] = \sum_{i=1}^{N-1} a_i^2.$$

  Need to understand the values a_i.
- Partition $\{0,1\}^t$ into certain prefix blocks $\{B_x : x \in \mathcal{L}\}$.
- Partition $[N]$ into discrete intervals $I(x)$ of strings with prefix x.

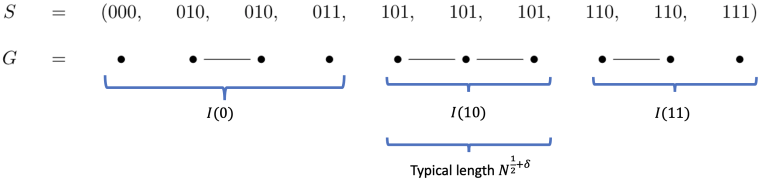


- Arrange that $\mathbb{E}[|I(x)|] \approx N^{\frac{1}{2}+\delta}$ for each $x \in \mathcal{L}$.

- How to understand $\mathbb{E}[|E(G,G')|]$? With $a_i = \mathbb{P}[(i,i+1) \in E(G)]$,

$$\mathbb{E}[|E(G,G')|] = \sum_{i=1}^{N-1} a_i^2 .$$

  Need to understand the values a_i.
- Partition $\{0,1\}^t$ into certain prefix blocks $\{B_x : x \in \mathcal{L}\}$.
- Partition $[N]$ into discrete intervals $I(x)$ of strings with prefix x.



- Arrange that $\mathbb{E}[|I(x)|] \approx N^{\frac{1}{2}+\delta}$ for each $x \in \mathcal{L}$.
- Key is **local homogeneity**: edge probability a_i acts constant on each $I(x)$.

- Typically, $|I(x)| \approx N^{\frac{1}{2}+\delta}$. Boundary fluctuations have smaller scale $N^{\frac{1}{2}}$.

Fluctuations of $I(x)$

- Typically, $|I(x)| \approx N^{\frac{1}{2}+\delta}$. Boundary fluctuations have smaller scale $N^{\frac{1}{2}}$.
- Hence $I(x)$ are almost deterministic. IID samples look like:

Fluctuations of $I(x)$

- Typically, $|I(x)| \approx N^{\frac{1}{2}+\delta}$. Boundary fluctuations have smaller scale $N^{\frac{1}{2}}$.

- Hence $I(x)$ are almost deterministic. IID samples look like:

- Leads to sum-over-blocks estimate

$$\mathbb{E}[|E(G,G')|] \lesssim \mathbb{E}\sum_{x\in\mathcal{L}} |E(G_{B_x}, G'_{B_x})|.$$

Fluctuations of $I(x)$

- Typically, $|I(x)| \approx N^{\frac{1}{2}+\delta}$. Boundary fluctuations have smaller scale $N^{\frac{1}{2}}$.

- Hence $I(x)$ are almost deterministic. IID samples look like:

- Leads to sum-over-blocks estimate

$$\mathbb{E}[|E(G, G')|] \lesssim \mathbb{E} \sum_{x \in \mathcal{L}} |E(G_{B_x}, G'_{B_x})|.$$

- Conversely: boundary fluctuation size $N^{\frac{1}{2}}$ is **almost** $|I(x)| \approx N^{\frac{1}{2}+\delta}$.
- These fluctuations act as convolutions to **locally homogenize** a_i.

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.
- Moreover, e_x depends only on the **digit profile** of x.

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.
- Moreover, e_x depends only on the **digit profile** of x.
  - Digit profile (c_0, c_1) means x contains $c_0 \log N$ digits 0 and $c_1 \log N$ digits 1.
- Total number of digit profiles is small: just $\log(N)^2 \ll N^{o(1)}$.

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.
- Moreover, e_x depends only on the **digit profile** of x.
  - Digit profile (c_0, c_1) means x contains $c_0 \log N$ digits 0 and $c_1 \log N$ digits 1.
- Total number of digit profiles is small: just $\log(N)^2 \ll N^{o(1)}$.
- Remains to find the digit profile with largest contribution:
  1. Count prefixes x with digit profile (c_0, c_1).

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.
- Moreover, e_x depends only on the **digit profile** of x.
  - Digit profile (c_0, c_1) means x contains $c_0 \log N$ digits 0 and $c_1 \log N$ digits 1.

- Total number of digit profiles is small: just $\log(N)^2 \ll N^{o(1)}$.

- Remains to find the digit profile with largest contribution:
  1. Count prefixes x with digit profile (c_0, c_1).
  2. Multiply by $E(G_{B_x}, G'_{B_x}) \approx N^{e_x}$ for total contribution from (c_0, c_1).

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.
- Moreover, e_x depends only on the **digit profile** of x.
  - Digit profile (c_0, c_1) means x contains $c_0 \log N$ digits 0 and $c_1 \log N$ digits 1.

- Total number of digit profiles is small: just $\log(N)^2 \ll N^{o(1)}$.

- Remains to find the digit profile with largest contribution:
  1. Count prefixes x with digit profile (c_0, c_1).
  2. Multiply by $E(G_{B_x}, G'_{B_x}) \approx N^{e_x}$ for total contribution from (c_0, c_1).
  3. Find the profile (c_0, c_1) with maximal contribution N^{α_*}.

Maximizing Over Digit Profiles

- Result: can explicitly compute $\mathbb{E}[|E(G_{B_x}, G'_{B_x})|] \approx N^{e_x}$.
- Moreover, e_x depends only on the **digit profile** of x.
  - Digit profile (c_0, c_1) means x contains $c_0 \log N$ digits 0 and $c_1 \log N$ digits 1.

- Total number of digit profiles is small: just $\log(N)^2 \ll N^{o(1)}$.

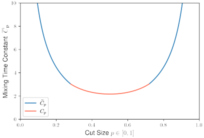
- Remains to find the digit profile with largest contribution:
  1. Count prefixes x with digit profile (c_0, c_1).
  2. Multiply by $E(G_{B_x}, G'_{B_x}) \approx N^{e_x}$ for total contribution from (c_0, c_1).
  3. Find the profile (c_0, c_1) with maximal contribution N^{α_*}.
- Maximum exponent α_* occurs at

$$(c_0^*, c_1^*) \sim \left(\frac{p^{\theta_p}}{p^{\theta_p} + q^{\theta_p}}, \frac{q^{\theta_p}}{p^{\theta_p} + q^{\theta_p}} \right).$$

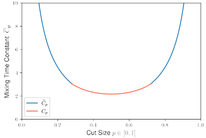
where $p^{\theta_p} + q^{\theta_p} = (p^2 + q^2)^2$. Leads to the threshold $C_p = \frac{3+\theta_p}{4\log(1/(p^2+q^2))}$.

Finishing the Upper Bound

- The derivation so far didn't suggest any phase transition...



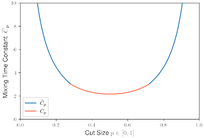
- The derivation so far didn't suggest any phase transition...



- Last issue: the discussion above assumed $N^{\frac{1}{2}}$ fluctuations of order statistics. This is wrong near the edge for prefixes $000\cdots$ or $111\cdots$.
  - Reason: $Var(Bin(N, r)) = Nr(1 - r) \ll N$ if $r \approx 0$ or $r \approx 1$.

Finishing the Upper Bound

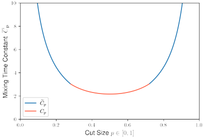
- The derivation so far didn't suggest any phase transition...



- Last issue: the discussion above assumed $N^{\frac{1}{2}}$ fluctuations of order statistics. This is wrong near the edge for prefixes $000\cdots$ or $111\cdots$.
  - Reason: $Var(Bin(N,r)) = Nr(1-r) \ll N$ if $r \approx 0$ or $r \approx 1$.

- The other extremes $x = 000\cdots 0$ and $x = 111\cdots 1$ yield another threshold $\widetilde{C}_p = \frac{1}{\log(1/\max(p,1-p))}$. "Closeness to the edge" has a linear effect on the exponent, so only extremes matter.

- The derivation so far didn't suggest any phase transition...



- Last issue: the discussion above assumed $N^{\frac{1}{2}}$ fluctuations of order statistics. This is wrong near the edge for prefixes $000\cdots$ or $111\cdots$.
  - Reason: $Var(Bin(N,r)) = Nr(1-r) \ll N$ if $r \approx 0$ or $r \approx 1$.

- The other extremes $x = 000\cdots0$ and $x = 111\cdots1$ yield another threshold $\widetilde{C}_p = \frac{1}{\log(1/\max(p,1-p))}$. "Closeness to the edge" has a linear effect on the exponent, so only extremes matter.

- Combining shows the desired upper bound:

$$t_{\mathrm{mix}} \le (\overline{C}_p + \varepsilon)\log N, \qquad \overline{C}_p = \max(C_p, \widetilde{C}_p).$$

Lower Bound: Returning to Cold Spots

1. Show mixing if $|E(G,G')|$ has small truncated exponential moments:

$$\mathbb{E}[e^{c|E(G,G')|}] \leq 1 + N^{-\delta}.$$

2. Reduce exponential moment estimate to first moment control

$$\mathbb{E}[|E(G,G')|] \leq N^{-\delta}.$$

3. Understand first moment $\implies$ upper bound t_{mix}.

- Main contribution to first moment $\implies$ optimal choice of cold spots to lower bound t_{mix}.

- Here we want to distinguish the distribution of π^G from uniform.

Lower-Bounding the Mixing Time

- Here we want to distinguish the distribution of π^G from uniform.
- Cold spots idea of [Lalley 2000]: construct non-random set $H \subseteq [N]$ typically containing all strings with "optimal" prefix digit profile (c_0^*, c_1^*).

- Here we want to distinguish the distribution of π^G from uniform.

- Cold spots idea of [Lalley 2000]: construct non-random set $H \subseteq [N]$ typically containing all strings with "optimal" prefix digit profile (c_0^*, c_1^*).
- These strings contribute $\gg |H|^{\frac{1}{2}+\delta}$ G-edges, all inside H.
- Each G-edge contributes ~ 1 ascent to π^G.

- Here we want to distinguish the distribution of π^G from uniform.

- Cold spots idea of [Lalley 2000]: construct non-random set $H \subseteq [N]$ typically containing all strings with "optimal" prefix digit profile (c_0^*, c_1^*).
- These strings contribute $\gg |H|^{\frac{1}{2}+\delta}$ G-edges, all inside H.
- Each G-edge contributes ~ 1 ascent to π^G.

- For uniform permutations, [#ascents in H] has $O(|H|^{1/2})$ fluctuations.

- Here we want to distinguish the distribution of π^G from uniform.

- Cold spots idea of [Lalley 2000]: construct non-random set $H \subseteq [N]$ typically containing all strings with "optimal" prefix digit profile (c_0^*, c_1^*).
- These strings contribute $\gg |H|^{\frac{1}{2}+\delta}$ G-edges, all inside H.
- Each G-edge contributes ~ 1 ascent to π^G.

- For uniform permutations, [#ascents in H] has $O(|H|^{1/2})$ fluctuations.
- Therefore, [#ascents in H] distinguishes π vs π^G.

Lower-Bounding the Mixing Time

- Here we want to distinguish the distribution of π^G from uniform.

- Cold spots idea of [Lalley 2000]: construct non-random set $H \subseteq [N]$ typically containing all strings with "optimal" prefix digit profile (c_0^*, c_1^*).
- These strings contribute $\gg |H|^{\frac{1}{2}+\delta}$ G-edges, all inside H.
- Each G-edge contributes ~ 1 ascent to π^G.

- For uniform permutations, [#ascents in H] has $O(|H|^{1/2})$ fluctuations.
- Therefore, [#ascents in H] distinguishes π vs π^G.

- Some work is needed to control the number of G-edges within H. [Lalley 2000] found 1st and 2nd moments, which only suffices for $p \approx 1/2$.

- Here we want to distinguish the distribution of π^G from uniform.

- Cold spots idea of [Lalley 2000]: construct non-random set $H \subseteq [N]$ typically containing all strings with "optimal" prefix digit profile (c_0^*, c_1^*).
- These strings contribute $\gg |H|^{\frac{1}{2}+\delta}$ G-edges, all inside H.
- Each G-edge contributes ~ 1 ascent to π^G.

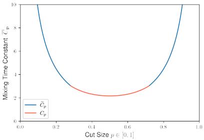
- For uniform permutations, [#ascents in H] has $O(|H|^{1/2})$ fluctuations.
- Therefore, [#ascents in H] distinguishes π vs π^G.

- Some work is needed to control the number of G-edges within H. [Lalley 2000] found 1st and 2nd moments, which only suffices for $p \approx 1/2$.
- For general p, truncate again – restrict also the **suffix** digit distribution.

Recap

- Main result: for $p \in (0,1)$, p-biased riffle shuffle exhibits cutoff at

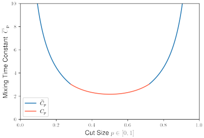
$$t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log(N).$$



Recap

- Main result: for $p \in (0,1)$, p-biased riffle shuffle exhibits cutoff at

$$t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log(N).$$

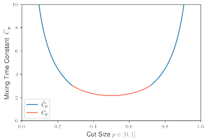


- Asymmetry breaks classical [Bayer-Diaconis 92] rising sequence analysis.

Recap

- Main result: for $p \in (0,1)$, p-biased riffle shuffle exhibits cutoff at

$$t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log(N).$$



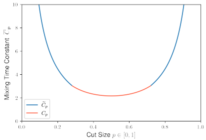
- Asymmetry breaks classical [Bayer-Diaconis 92] rising sequence analysis.

- First step: consider transformed problem involving strings $(s_1, \ldots, s_N)$, associated "shuffle graphs" G, and transformed permutations $\pi \to \pi^G$.

Recap

- Main result: for $p \in (0,1)$, p-biased riffle shuffle exhibits cutoff at

$$t_{\mathrm{mix}} = (\overline{C}_p \pm o(1))\log(N).$$



- Asymmetry breaks classical [Bayer-Diaconis 92] rising sequence analysis.

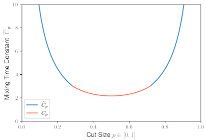
- First step: consider transformed problem involving strings $(s_1, \ldots, s_N)$, associated "shuffle graphs" G, and transformed permutations $\pi \to \pi^G$.

- Key quantity: edge intersection $|E(G, G')|$ of independent copies G, G'. Show mixing by bounding the exponential moment.

Recap

- Main result: for $p \in (0,1)$, p-biased riffle shuffle exhibits cutoff at

$$t_{\mathrm{mix}} = (\overline{C}_p \pm o(1)) \log(N).$$



- Asymmetry breaks classical [Bayer-Diaconis 92] rising sequence analysis.

- First step: consider transformed problem involving strings $(s_1, \ldots, s_N)$, associated "shuffle graphs" G, and transformed permutations $\pi \to \pi^G$.

- Key quantity: edge intersection $|E(G, G')|$ of independent copies G, G'. Show mixing by bounding the exponential moment.

- Main obstruction to mixing: **cold spots** with many G-edges $\implies$ many ascents in the inverse shuffle permutation π^G.