Characterization of cutoff for reversible Markov chains

Yuval Peres

Joint work with Riddhi Basu and Jonathan Hermon

3 December 2014
Transition matrix - $P$ (reversible).

Stationary dist. - $\pi$.

Reversibility: $\pi(x)P(x, y) = \pi(y)P(y, x)$, $\forall x, y \in \Omega$.

Laziness $P(x, x) \geq 1/2$, $\forall x \in \Omega$. 

Joint work with Riddhi Basu and Jonathan Hermon
For any 2 dist. \( \mu, \nu \) on \( \Omega \), their total-variation distance is:

\[
\| \mu - \nu \|_{TV} \overset{d}{=} \max_{A \subset \Omega} \mu(A) - \nu(A).
\]

\[
d(t, x) \overset{d}{=} \| P^t_x - \pi \|_{TV}, \quad d(t) \overset{d}{=} \max_{x \in \Omega} d(t, x).
\]
TV distance

For any 2 dist. \( \mu, \nu \) on \( \Omega \), their **total-variation distance** is:

\[
\| \mu - \nu \|_{TV} = \max_{A \subset \Omega} \mu(A) - \nu(A).
\]

\[
d(t, x) \overset{d}{=} \| P_t x - \pi \|_{TV}, \quad d(t) \overset{d}{=} \max_{x \in \Omega} d(t, x).
\]

The **\( \epsilon \)-mixing-time** \( (0 < \epsilon < 1) \) is:

\[
t_{\text{mix}}(\epsilon) \overset{d}{=} \min \{ t : d(t) \leq \epsilon \}
\]

\[
t_{\text{mix}} \overset{d}{=} t_{\text{mix}}(1/4).
\]
Cutoff - definition

Def: a sequence of MCs \( (X_t^{(n)}) \) exhibits \textbf{cutoff} if

\[
t_{\text{mix}}^{(n)}(\epsilon) - t_{\text{mix}}^{(n)}(1 - \epsilon) = o(t_{\text{mix}}^{(n)}), \forall 0 < \epsilon < 1/4.
\]  

(1)
Cutoff - definition

- Def: a sequence of MCs \( (X_t^{(n)}) \) exhibits \textbf{cutoff} if

\[
t^{(n)}_\text{mix}(\epsilon) - t^{(n)}_\text{mix}(1 - \epsilon) = o(t^{(n)}_\text{mix}), \quad \forall \; 0 < \epsilon < 1/4. \tag{1}
\]

- \((w_n)\) is called a \textbf{cutoff window} for \((X_t^{(n)})\) if: \(w_n = o\left(t^{(n)}_\text{mix}\right)\), and

\[
t^{(n)}_\text{mix}(\epsilon) - t^{(n)}_\text{mix}(1 - \epsilon) \leq c\epsilon w_n, \quad \forall n \geq 1, \forall \epsilon \in (0, 1/4).
\]
Cutoff

Figure: cutoff
Cutoff was first identified for random transpositions Diaconis & Shashahani 81 and RW on the hypercube by Aldous 83.
Cutoff was first identified for random transpositions Diaconis & Shashahani 81 and RW on the hypercube by Aldous 83.

Many chains are believed to exhibit cutoff. Verifying the occurrence of cutoff rigorously is usually hard.
Cutoff was first identified for random transpositions Diaconis & Shashahani 81 and RW on the hypercube by Aldous 83.

Many chains are believed to exhibit cutoff. Verifying the occurrence of cutoff rigorously is usually hard.

The name cutoff was coined by Aldous and Diaconis in their seminal 86 paper.

Aldous & Diaconis 86 - “the most interesting open problem”: Find verifiable conditions for cutoff.
Cutoff was first identified for random transpositions Diaconis & Shashahani 81 and RW on the hypercube by Aldous 83.

Many chains are believed to exhibit cutoff. Verifying the occurrence of cutoff rigorously is usually hard.

The name cutoff was coined by Aldous and Diaconis in their seminal 86 paper.

Aldous & Diaconis 86 - “the most interesting open problem”: Find verifiable conditions for cutoff.
Let $\lambda_2$ be the largest non-trivial e.v. of $P$.

Definition: $\text{gap} = 1 - \lambda_2$ - the **spectral gap**.

Def: $t_{\text{rel}} := \text{gap}^{-1}$ - the **relaxation-time**.
The product condition (Prod. cond.)

- In a 2004 Aim workshop I proposed that **The product condition (Prod. Cond.)** -
  \( \text{gap}^{(n)} t^{(n)}_{\text{mix}} \to \infty \) (equivalently, \( t^{(n)}_{\text{rel}} = o(t^{(n)}_{\text{mix}}) \))
  should imply cutoff for "nice" reversible chains.
- (It is a necessary condition for cutoff)

Joint work with Riddhi Basu and Jonathan Hermon
In a 2004 Aim workshop I proposed that \textbf{The product condition (Prod. Cond.)} -
\[ \text{gap}^{(n)} t_{\text{mix}}^{(n)} \to \infty \] (equivalently, \( t_{\text{rel}}^{(n)} = o(t_{\text{mix}}^{(n)}) \)) should imply cutoff for "nice" reversible chains.

(It is a necessary condition for cutoff)

It is not always sufficient - examples due to Aldous and Pak.

Problem: Find families of MCs s.t. \textbf{Prod. Cond.} \( \implies \) cutoff.
Aldous’ example

- The mass is concentrated in a small neighborhood of $y$.
- Last step away from $z$ before $T_y$ “determines” $T_y$.

Figure: Fixed bias to the right conditioned on a non-lazy step.

- Different laziness probabilities along the 2 paths.
- $t_{\text{rel}}^{(n)} = O(1)$.
- $d_n(t) \sim P_x[T_y > t] \implies \epsilon \leq d_n(130n) \leq d_n(128n) \leq 1 - \epsilon$, for some $\epsilon$. 

Joint work with Riddhi Basu and Jonathan Hermon
Characterization of cutoff for reversible Markov chains
Aldous’ example

Figure: Fixed bias to the right conditioned on a non-lazy step.

$d_n(t)$

126n

132n
Def: The hitting time of a set $A \subset \Omega = T_A := \min\{t : X_t \in A\}$.

Remark: (2) may fail for $\alpha > \frac{1}{2}$. Joint work with Riddhi Basu and Jonathan Hermon.
Hitting and Mixing

- **Def:** The **hitting time** of a set $A \subset \Omega = T_A := \min\{t : X_t \in A\}$.

- Hitting times of “worst” sets are related to mixing - mid 80’s (Aldous).

- Refined independently by Oliviera (2011) and Peres-Sousi (2011) (case $\alpha = 1/2$ due to Griffiths-Kang-Oliviera-Patel 2012): for any irreducible reversible lazy MC and $0 < \alpha \leq 1/2$:

  $$t_H(\alpha) = \Theta_\alpha(t_{\text{mix}}), \text{ where } t_H(\alpha) := \max_{x,A: \pi(A) \geq \alpha} \mathbb{E}_x[T_A].$$  

  (2)

Remark: (2) may fail for $\alpha > 1/2$. 

Joint work with Riddhi Basu and Jonathan Hermon
Hitting and Mixing

- **Def:** The **hitting time** of a set $A \subset \Omega = T_A := \min\{t : X_t \in A\}$.

- Hitting times of “worst” sets are related to mixing - mid 80’s (Aldous).

- Refined independently by Oliviera (2011) and Peres-Sousi (2011) (case $\alpha = 1/2$ due to Griffiths-Kang-Oliviera-Patel 2012): for any irreducible reversible lazy MC and $0 < \alpha \leq 1/2$:

  \[ t_H(\alpha) = \Theta(\alpha(t_{\text{mix}})), \text{ where } t_H(\alpha) := \max_{x, A : \pi(A) \geq \alpha} \mathbb{E}_x[T_A]. \quad (2) \]

- We relate $d(t)$ and $\max_{x, A : \pi(A) \geq \alpha} \mathbb{P}_x[T_A > t]$ and refine (2) by also allowing $1/2 < \alpha \leq 1 - \exp[-Ct_{\text{mix}}/t_{\text{rel}}]$ and improving $\Theta(\alpha)$ to $\Theta$.

- **Remark:** (2) may fail for $\alpha > 1/2$.
counter-example

Figure: $n$ is the index of the chain
Concentration of hitting times of “worst” sets is related to cutoff in birth and death (BD) chains.
Concentration of hitting times of “worst” sets is related to cutoff in birth and death (BD) chains.

Diaconis & Saloff-Coste (06) (separation cutoff) and Ding-Lubetzky-Peres (10) (TV cutoff):

A seq. of BD chains exhibits cutoff iff the Prod. Cond. holds.
Concentration of hitting times of “worst” sets is related to cutoff in birth and death (BD) chains.

Diaconis & Saloff-Coste (06) (separation cutoff) and Ding-Lubetzky-Peres (10) (TV cutoff):

A seq. of BD chains exhibits cutoff iff the Prod. Cond. holds.

We extend their results to weighted nearest-neighbor RWs on trees.
Cutoff for trees

**Theorem**

Let \( (V, P, \pi) \) be a lazy Markov chain on a tree \( T = (V, E) \) with \( |V| \geq 3 \). Then

\[
    t_{\text{mix}}(\epsilon) - t_{\text{mix}}(1 - \epsilon) \leq C \sqrt{|\log \epsilon| t_{\text{rel}} t_{\text{mix}}}, \text{ for any } 0 < \epsilon \leq 1/4.
\]

*In particular, the Prod. Cond. implies cutoff with a cutoff window \( w_n = \sqrt{t_{\text{rel}}(n) t_{\text{mix}}(n)} \) and*

\[
    c_\epsilon = C \sqrt{|\log \epsilon|}.
\]
Cutoff for trees

Theorem

Let \((V, P, \pi)\) be a lazy Markov chain on a tree \(T = (V, E)\) with \(|V| \geq 3\). Then

\[
t_{\text{mix}}(\epsilon) - t_{\text{mix}}(1 - \epsilon) \leq C \sqrt{|\log \epsilon| t_{\text{rel}} t_{\text{mix}}}, \text{ for any } 0 < \epsilon \leq 1/4.
\]

In particular, the Prod. Cond. implies cutoff with a cutoff window \(w_n = \sqrt{t_{(n)}^{(n)} t_{\text{rel}} t_{\text{mix}}}\) and

\[
c_{\epsilon} = C \sqrt{|\log \epsilon|}.
\]

- Ding Lubetzky Peres (10) - For BD chains \(t_{\text{mix}}(\epsilon) - t_{\text{mix}}(1 - \epsilon) \leq O(\epsilon^{-1} \sqrt{t_{\text{rel}} t_{\text{mix}}})\)

and in some cases \(w_n = \Omega \left(\sqrt{t_{(n)}^{(n)} t_{\text{rel}} t_{\text{mix}}}\right)\).
To mix - escape and then relax

- Definition: $\text{hit}_\alpha := \text{hit}_\alpha(1/4)$, where

$$
\text{hit}_{\alpha,x}(\epsilon) := \min \{ t : P_x[T_A > t] \leq \epsilon : \text{for all } A \subset \Omega \text{ s.t. } \pi(A) \geq \alpha \},
$$

$$
\text{hit}_\alpha(\epsilon) := \max_{x \in \Omega} \text{hit}_{\alpha,x}(\epsilon)
$$
To mix - escape and then relax

- **Definition:** \( \text{hit}_\alpha := \text{hit}_\alpha (1/4) \), where

\[
\text{hit}_{\alpha, x}(\epsilon) := \min \{ t : P_x[T_A > t] \leq \epsilon : \text{for all } A \subset \Omega \text{ s.t. } \pi(A) \geq \alpha \},
\]

\[
\text{hit}_\alpha(\epsilon) := \max_{x \in \Omega} \text{hit}_{\alpha, x}(\epsilon)
\]

- **Easy direction:** to mix, the chain must first escape from small sets = “first stage of mixing”.

Joint work with Riddhi Basu and Jonathan Hermon

Characterization of cutoff for reversible Markov chains
To mix - escape and then relax

- Definition: \( \text{hit}_\alpha := \text{hit}_\alpha(1/4) \), where

\[
\text{hit}_{\alpha,x}(\epsilon) := \min\{t : P_x[T_A > t] \leq \epsilon : \text{for all } A \subset \Omega \text{ s.t. } \pi(A) \geq \alpha\},
\]

\[
\text{hit}_\alpha(\epsilon) := \max_{x \in \Omega} \text{hit}_{\alpha,x}(\epsilon)
\]

- Easy direction: to mix, the chain must first escape from small sets = “first stage of mixing”.

- Loosely speaking - we show that in the 2nd “stage of mixing”, the chain mixes at the fastest possible rate (governed by its relaxation-time).
Fact: Let $A \subset \Omega$ be such that $\pi(A) \geq 1/2$. Then (under reversibility)

$$
P_\pi[T_A > 2st_{rel}] \leq \frac{e^{-s}}{2}, \text{ for all } s \geq 0.
$$
Hitting times when \( X_0 \sim \pi \)

- Fact: Let \( A \subset \Omega \) be such that \( \pi(A) \geq 1/2 \). Then (under reversibility)

\[
P_\pi[T_A > 2st_{rel}] \leq \frac{e^{-s}}{2}, \text{ for all } s \geq 0.
\]

- By a coupling argument,

\[
P_x[T_A > t + 2st_{rel}] \leq d(t) + P_\pi[T_A > 2st_{rel}].
\]
Hitting of worst sets

For any reversible irreducible finite lazy chain and any $0 < \epsilon \leq 1/4$,

$$\text{hit}_{1/2}(3\epsilon) - t_{\text{rel}}|\log(2\epsilon)| \leq t_{\text{mix}}(2\epsilon) \leq \text{hit}_{1/2}(\epsilon) + t_{\text{rel}}|\log(4\epsilon)|$$

- Terms involving $t_{\text{rel}}$ are negligible under the Prod. Cond.
For any reversible irreducible finite lazy chain and any $0 < \epsilon \leq 1/4$,

\[
\text{hit}_{1/2}(3\epsilon) - t_{\text{rel}}|\log(2\epsilon)| \leq t_{\text{mix}}(2\epsilon) \leq \text{hit}_{1/2}(\epsilon) + t_{\text{rel}}|\log(4\epsilon)|
\]

- Terms involving $t_{\text{rel}}$ are negligible under the Prod. Cond.
For any reversible irreducible finite lazy chain and any $0 < \epsilon \leq 1/4$, 

\[ \text{hit}_{1/2}(3\epsilon) - t_{\text{rel}}|\log(2\epsilon)| \leq t_{\text{mix}}(2\epsilon) \leq \text{hit}_{1/2}(\epsilon) + t_{\text{rel}}|\log(4\epsilon)| \]

- Terms involving $t_{\text{rel}}$ are negligible under the Prod. Cond..
- A similar two sided inequality holds for $t_{\text{mix}}(1 - 2\epsilon)$. 

Joint work with Riddhi Basu and Jonathan Hermon
Main abstract result

Definition: A sequence has $\text{hit}_\alpha$-cutoff if

$$\text{hit}_\alpha^{(n)}(\epsilon) - \text{hit}_\alpha^{(n)}(1 - \epsilon) = o(\text{hit}_\alpha^{(n)}) \text{ for all } 0 < \epsilon < 1/4.$$
Main abstract result

Definition: A sequence has $\text{hit}_\alpha$-cutoff if

$$\text{hit}_\alpha^n(\epsilon) - \text{hit}_\alpha^n(1 - \epsilon) = o(\text{hit}_\alpha^n) \text{ for all } 0 < \epsilon < 1/4.$$ 

Main abstract result:

Theorem

Let $(\Omega_n, P_n, \pi_n)$ be a seq. of finite reversible lazy MCs. Then TFAE:

- The seq. exhibits cutoff.
- The seq. exhibits a $\text{hit}_\alpha$-cutoff for some $\alpha \in (0, 1/2)$.
- The seq. exhibits a $\text{hit}_\alpha$-cutoff for some $\alpha \in [1/2, 1)$ and the Prod. Cond. holds.
Main abstract result

Definition: A sequence has $\text{hit}_\alpha$-cutoff if

$$\text{hit}_\alpha(n)(\epsilon) - \text{hit}_\alpha(n)(1 - \epsilon) = o(\text{hit}_\alpha(n)) \text{ for all } 0 < \epsilon < 1/4.$$ 

Main abstract result:

Theorem

Let $(\Omega_n, P_n, \pi_n)$ be a seq. of finite reversible lazy MCs. Then TFAE:

- The seq. exhibits cutoff.
- The seq. exhibits a $\text{hit}_\alpha$-cutoff for some $\alpha \in (0, 1/2)$.
- The seq. exhibits a $\text{hit}_\alpha$-cutoff for some $\alpha \in [1/2, 1)$ and the Prod. Cond. holds.

The equivalence of cutoff to $\text{hit}_{1/2}$-cutoff under the Prod. Cond. follows from the ineq. from the prev. slide together with the fact that $\text{hit}_{1/2} = \Theta(t_{\text{mix}}(n))$. 

Joint work with Riddhi Basu and Jonathan Hermon
Main abstract result

Definition: A sequence has $\text{hit}_\alpha$-cutoff if

$$\text{hit}_\alpha^{(n)}(\epsilon) - \text{hit}_\alpha^{(n)}(1 - \epsilon) = o(\text{hit}_\alpha^{(n)}) \text{ for all } 0 < \epsilon < 1/4.$$ 

Main abstract result:

**Theorem**

Let $(\Omega_n, P_n, \pi_n)$ be a seq. of finite reversible lazy MCs. Then TFAE:

- The seq. exhibits cutoff.
- The seq. exhibits a $\text{hit}_\alpha$-cutoff for some $\alpha \in (0, 1/2)$.
- The seq. exhibits a $\text{hit}_\alpha$-cutoff for some $\alpha \in [1/2, 1)$ and the Prod. Cond. holds.

The equivalence of cutoff to $\text{hit}_{1/2}$-cutoff under the Prod. Cond. follows from the ineq. from the prev. slide together with the fact that $\text{hit}_{1/2}^{(n)} = \Theta(t_{\text{mix}}^{(n)})$.

For general $\alpha$ we show under the Prod. Cond. (using the tail decay of $T_A/t_{\text{rel}}$ when $X_0 \sim \pi$):

$$\text{hit}_\alpha\text{-cutoff for some } \alpha \in (0, 1) \implies \text{hit}_\beta\text{-cutoff for all } \beta \in (0, 1).$$

Joint work with Riddhi Basu and Jonathan Hermon
Def: For $f \in \mathbb{R}^\Omega$, $t \geq 0$, define $P^t f \in \mathbb{R}^\Omega$ by

$$P^t f(x) := \mathbb{E}_x[f(X_t)] = \sum_y P^t(x, y) f(y).$$
Def: For \( f \in \mathbb{R}^\Omega, t \geq 0 \), define \( P^t f \in \mathbb{R}^\Omega \) by
\[
P^t f(x) := \mathbb{E}_x[f(X_t)] = \sum_y P^t(x, y) f(y).
\]

For \( f \in \mathbb{R}^\Omega \) define \( \mathbb{E}_\pi[f] := \sum_{x \in \Omega} \pi(x) f(x) \) and \( \|f\|_2^2 := \mathbb{E}_\pi[f^2]. \)
Def: For $f \in \mathbb{R}^\Omega$, $t \geq 0$, define $P^t f \in \mathbb{R}^\Omega$ by

$$P^t f(x) := \mathbb{E}_x[f(X_t)] = \sum_y P^t(x, y)f(y).$$

For $f \in \mathbb{R}^\Omega$ define $\mathbb{E}_\pi[f] := \sum_{x \in \Omega} \pi(x)f(x)$ and $\|f\|_2^2 := \mathbb{E}_\pi[f^2]$.

For $g \in \mathbb{R}^\Omega$ denote $\text{Var}_\pi g := \|g - \mathbb{E}_\pi[g]\|_2^2$. 

The following is well-known and follows from elementary linear-algebra.

**Lemma (Contraction Lemma)**

Let $(\Omega, P, \pi)$ be a finite rev. irr. lazy MC. Let $A \subset \Omega$. Let $t \geq 0$. Then

$$\text{Var}_\pi P^t 1_A \leq e^{-t/\text{rel}}.$$

(3)
Def: For $f \in \mathbb{R}^{\Omega}$, $t \geq 0$, define $P^t f \in \mathbb{R}^{\Omega}$ by

$$P^t f(x) := \mathbb{E}_x[f(X_t)] = \sum_y P^t(x,y) f(y).$$

For $f \in \mathbb{R}^{\Omega}$ define $\mathbb{E}_\pi[f] := \sum_{x \in \Omega} \pi(x) f(x)$ and $\|f\|_2^2 := \mathbb{E}_\pi[f^2]$.

For $g \in \mathbb{R}^{\Omega}$ denote $\text{Var}_\pi g := \|g - \mathbb{E}_\pi[g]\|_2^2$.

The following is well-known and follows from elementary linear-algebra.

**Lemma (Contraction Lemma)**

*Let $(\Omega, P, \pi)$ be a finite rev. irr. lazy MC. Let $A \subset \Omega$. Let $t \geq 0$. Then*

$$\text{Var}_\pi P^t 1_A \leq e^{-2t/t_{rel}}. \tag{3}$$
Maximal Inequality

The main ingredient in our approach is Starr’s maximal-inequality (66) (refines Stein’s max-inequality (61))

Theorem (Maximal inequality)

Let \((\Omega, P, \pi)\) be a lazy irreducible reversible Markov chain. Let \(f \in \mathbb{R}^\Omega\). Define the corresponding maximal function \(f^* \in \mathbb{R}^\Omega\) as

\[
f^*(x) := \sup_{0 \leq k < \infty} |P^k(f)(x)| = \sup_{0 \leq k < \infty} |\mathbb{E}_x[f(X_k)]|.
\]

Then for \(1 < p < \infty\),

\[
\|f^*\|_p \leq q\|f\|_p \quad 1/p + 1/q = 1
\]

(4)

Joint work with Riddhi Basu and Jonathan Hermon

Characterization of cutoff for reversible Markov chains
Combining the Max-in. with the Contraction Lemma

Goal: want for every $A \subset \Omega$ to have $G = G_s(A) \subset \Omega$ s.t. $T_G \leq t$ serves as a certificate of “being $\epsilon$-mixed w.r.t. $A$” and to control its $\pi$ measure from below.
Combining the Max-in. with the Contraction Lemma

Goal: want for every $A \subset \Omega$ to have $G = G_s(A) \subset \Omega$ s.t. $T_G \leq t$ serves as a certificate of “being $\epsilon$-mixed w.r.t. $A$” and to control its $\pi$ measure from below.

- Let $\sigma_s := e^{-s/t_{rel}} \geq \sqrt{\text{Var}_\pi P_s 1_A}$ (contraction lemma).
- Consider

$$G = G_s(A) := \left\{ g : \forall \tilde{s} \geq s, |P_{\tilde{s}}^g(A) - \pi(A)| \leq 4\sigma_s \right\}.$$
Combining the Max-in. with the Contraction Lemma

Goal: want for every $A \subset \Omega$ to have $G = G_s(A) \subset \Omega$ s.t. $T_G \leq t$ serves as a certificate of “being $\epsilon$-mixed w.r.t. $A$” and to control its $\pi$ measure from below.

- Let $\sigma_s := e^{-s/t_{rel}} \geq \sqrt{\text{Var}_\pi P^s 1_A}$ (contraction lemma).
- Consider
  \[
  G = G_s(A) := \left\{ g : \forall \tilde{s} \geq s, |P^\tilde{s}_g(A) - \pi(A)| \leq 4\sigma_s \right\}.
  \]

- Want precision $4\sigma_s = \epsilon \implies s := t_{rel} \times \log(4/\epsilon)$.
Combining the Max-in. with the Contraction Lemma

Goal: want for every $A \subset \Omega$ to have $G = G_s(A) \subset \Omega$ s.t. $T_G \leq t$ serves as a certificate of “being $\epsilon$-mixed w.r.t. $A$” and to control its $\pi$ measure from below.

- Let $\sigma_s := e^{-s/t_{rel}} \geq \sqrt{\text{Var}_\pi P^1_A}$ (contraction lemma).
- Consider

$$G = G_s(A) := \{ g : \forall \tilde{s} \geq s, |P^{\tilde{s}}_g(A) - \pi(A)| \leq 4\sigma_s \}.$$

- Want precision $4\sigma_s = \epsilon \implies s := t_{rel} \times \log(4/\epsilon)$.

Claim

$$\pi(G) \geq 1/2. \quad (5)$$
Combining the Max-in. with the Contraction Lemma

Goal: want for every $A \subset \Omega$ to have $G = G_s(A) \subset \Omega$ s.t. $T_G \leq t$ serves as a certificate of “being $\epsilon$-mixed w.r.t. $A$” and to control its $\pi$ measure from below.

- Let $\sigma_s := e^{-s/t_{rel}} \geq \sqrt{\text{Var}_\pi P^s 1_A}$ (contraction lemma).
- Consider
  \[ G = G_s(A) := \{ g : \forall \tilde{s} \geq s, |P^\tilde{s}_g(A) - \pi(A)| \leq 4\sigma_s \} \, . \]

- Want precision $4\sigma_s = \epsilon \implies s := t_{rel} \times \log(4/\epsilon)$.

Claim

\[ \pi(G') \geq 1/2. \] (5)

Proof: Set $f_s := P^s (1_A - \pi(A))$. Then

\[ G^c \subset \{ f_s^* > 4\|f_s\|_2 \} \, . \]

Apply Starr’s inequality.

Joint work with Riddhi Basu and Jonathan Hermon
Claim:
\[ t_{\text{mix}}(2\epsilon) \leq \text{hit}_{1/2}(\epsilon) + t_{\text{rel}} \times \log(4/\epsilon). \]

**Proof:** Recall
\[ G := G_s(A, m) := \left\{ g : \forall \tilde{s} \geq s, |P_{\tilde{s}}(A) - \pi(A)| \leq \epsilon \right\}, \quad s := t_{\text{rel}} \times \log(4/\epsilon) \]

Set \( t := \text{hit}_{1/2}(\epsilon) \). By prev. claim \( \pi(G) \geq 1/2 \implies P_x[T_G > t] \leq \epsilon \) (by def. of \( t \)).
Main idea

Claim:

\[ t_{\text{mix}}(2\epsilon) \leq \text{hit}_{1/2}(\epsilon) + t_{\text{rel}} \times \log(4/\epsilon). \]

Proof: Recall

\[ G := G_s(A, m) := \left\{ g : \forall \tilde{s} \geq s, |P_{\tilde{s}}^g(A) - \pi(A)| \leq \epsilon \right\}, \quad s := t_{\text{rel}} \times \log(4/\epsilon) \]

Set \( t := \text{hit}_{1/2}(\epsilon). \) By prev. claim \( \pi(G) \geq 1/2 \implies P_x[T_G > t] \leq \epsilon \) (by def. of \( t \)).

For any \( x, A: \)

\[ |P_{x}^{t+s}(A) - \pi(A)| \leq P_x[T_G > t] + \max_{g \in G, \tilde{s} \geq s} |P_{\tilde{s}}^g(A) - \pi(A)| \leq 2\epsilon. \]
Let: \( T := (V, E) \) be a finite tree.

\((V, P, \pi)\) a lazy MC corresponding to some (lazy) weighted nearest-neighbor walk on \( T \) (i.e. \( P(x, y) > 0 \) iff \( \{x, y\} \in E \) or \( y = x \)).

Fact: (Kolmogorov’s cycle condition) every MC on a tree is reversible.
Can the tree structure be used to determine the identity of the “worst” sets?
Can the tree structure be used to determine the identity of the “worst” sets?

Easier question: what set of $\pi$ measure $\geq 1/2$ is the “hardest” to hit in a birth & death chain with state space $[n] := \{1, 2, \ldots, n\}$?
Can the tree structure be used to determine the identity of the “worst” sets?

Easier question: what set of $\pi$ measure $\geq 1/2$ is the “hardest” to hit in a birth & death chain with state space $[n] := \{1, 2, \ldots, n\}$?

Answer: take a state $m$ with $\pi([m]) \geq 1/2$ and $\pi([m - 1]) < 1/2$. Then the set worst set would be either $[m]$ or $[n] \setminus [m - 1]$. 

Joint work with Riddhi Basu and Jonathan Hermon
How to generalize this to trees?
Central vertex

Figure: A vertex $o \in V$ is called a central-vertex if each connected component of $T \setminus \{o\}$ has stationary probability at most $1/2$. 
There is always a central-vertex (and at most 2). We fix one, denote it by $o$ and call it the root.
There is always a central-vertex (and at most 2). We fix one, denote it by $o$ and call it the root.

It follows from our analysis that for trees the Prod. Cond. holds iff $T_o$ is concentrated (from worst leaf).

A counterintuitive result $\iff$ $\exists$ such unweighed trees (Peres-Sousi (13)).
Let $A$ be s.t. $\Pi(A) \geq 1/2$. Partition $V$ to $B$ and $D = V \setminus B$ s.t. $B$ is connected, $o$ is in $B$ and $\Pi(A') \geq 1/4$, where $A' := (D \cup \{o\}) \cap A$.

$$P_o[T_A > s] \leq P_o[T_{A'} > s] \leq P_{\Pi B}[T_{A'} > s] \leq 2P_{\Pi}[T_{A'} > s],$$

where $\Pi_B$ is $\Pi$ conditioned on $B$.

Take $s := C_{\text{rel}}|\log(\varepsilon)|$. Then $P_o[T_A > s] \leq \varepsilon$.

$$\Rightarrow \text{hit}_{1/2}(a+\varepsilon) \leq \min\{t: P_x[T_o > t] \leq a, \text{for all } x\} + s.$$

Trivially: $\min\{t: P_x[T_o > t] \leq a, \text{for all } x\} \leq \text{hit}_{1/2}(a)$.

Figure: Hitting the worst set is roughly like hitting $o$. 
Cutoff would follow if we show that $T_o$ is concentrated (under the Prod. Cond.).

More precisely, we need to show that $\mathbb{E}_x[T_o] = \Omega(t_{mix}) \implies T_{y^\beta}(x)$ is concentrated if $X_0 = x$. 

Joint work with Riddhi Basu and Jonathan Hermon
- Cutoff would follow if we show that $T_o$ is concentrated (under the Prod. Cond.).

- More precisely, we need to show that $\mathbb{E}_x[T_o] = \Omega(t_{\text{mix}}) \implies T_{y\beta}(x)$ is concentrated if $X_0 = x$. 

Joint work with Riddhi Basu and Jonathan Hermon
Characterization of cutoff for reversible Markov chains
Figure: Let $v_0 = x, v_1, \ldots, v_k = o$ be the vertices along the path from $x$ to $o$. 
Proof of Concentration: \( \text{Var}_x[T_o] \leq C t_{rel} t_{\text{mix}} \):

- It suffices to show that \( \text{Var}_x[T_o] \leq 4 t_{rel} \mathbb{E}_x[T_o] \).

Joint work with Riddhi Basu and Jonathan Hermon
Proof of Concentration: $\text{Var}_x[T_o] \leq C t_{\text{rel}} t_{\text{mix}}$:

- It suffices to show that $\text{Var}_x[T_o] \leq 4 t_{\text{rel}} \mathbb{E}_x[T_o]$.

- If $X_0 = x$ then $T_o$ is the sum of crossing times of the edges along the path between $x$: $\tau_i := T_{v_i} - T_{v_{i-1}} \overset{d}{=} T_{v_i} \text{ under } X_0 = v_{i-1}$
Trees

Proof of Concentration: $\text{Var}_x[T_o] \leq C t_{rel} t_{mix}$:

- It suffices to show that $\text{Var}_x[T_o] \leq 4 t_{rel} \mathbb{E}_x[T_o]$.

- If $X_0 = x$ then $T_o$ is the sum of crossing times of the edges along the path between $x$: $\tau_i := T_{v_i} - T_{v_{i-1}} \overset{d}{=} T_{v_i}$ under $X_0 = v_{i-1}$

- $\tau_1, \ldots, \tau_k$ are independent $\implies$ it suffices to bound the sum of their 2nd moments

$\text{Var}_x[T_o] = \sum \text{Var}_x[\tau_i] = \sum \text{Var}_{v_{i-1}}[T_{v_i}] \leq \sum \mathbb{E}_{v_{i-1}}[T_{v_i}^2]$. 

Proof of Concentration: $\text{Var}_x[T_o] \leq C t_{\text{rel}} t_{\text{mix}}$:

- It suffices to show that $\text{Var}_x[T_o] \leq 4 t_{\text{rel}} \mathbb{E}_x[T_o]$.

- If $X_0 = x$ then $T_o$ is the sum of crossing times of the edges along the path between $x$: $\tau_i := T_{v_i} - T_{v_{i-1}} \overset{d}{=} T_{v_i}$ under $X_0 = v_{i-1}$.

- $\tau_1, \ldots, \tau_k$ are independent $\implies$ it suffices to bound the sum of their 2nd moments
  
  $\text{Var}_x[T_o] = \sum \text{Var}_x[\tau_i] = \sum \text{Var}_{v_{i-1}}[T_{v_i}] \leq \sum \mathbb{E}_{v_{i-1}}[T_{v_i}^2]$.

- Denote the subtree rooted at $v$ (the set of vertices whose path to $o$ goes through $v$) by $W_v$. For $A \subset \Omega$ let $\pi_A$ be $\pi$ conditioned on $A$.

- Kac formula implies that for any $A$, there exists a dist. $\mu$ on the external vertex boundary of $A$ s.t. $\mathbb{E}_\mu[T_A^2] \leq 2 \mathbb{E}_\mu[T_A] \mathbb{E}_{\pi_A}(T_A) \implies$

- By the tree structure $\mathbb{E}_{v_{i-1}}[T_{v_i}^2] \leq 2 \mathbb{E}_{v_{i-1}}[T_{v_i}] \mathbb{E}_{\pi_{W_{v_{i-1}}}}[T_{v_i}]$.

- Not hard to show $\mathbb{E}_{\pi_{W_{v_{i-1}}}}[T_{v_i}] \leq 2 t_{\text{rel}}$ (generally $\pi(A) \mathbb{E}_{\pi_A}(T_A) \leq t_{\text{rel}}$) so

  \[
  \sum \mathbb{E}_{v_{i-1}}[T_{v_i}^2] \leq \sum 4 t_{\text{rel}} \mathbb{E}_{v_{i-1}}[T_{v_i}] = 4 t_{\text{rel}} \mathbb{E}_x[T_o].
  \]
Beyond trees

- The tree assumption can be relaxed. In particular, we can treat jumps to vertices of bounded distance on a tree (i.e. the length of the path from $u$ to $v$ in the tree (which is now just an auxiliary structure) is $> r \implies P(u, v) = 0$) under some mild necessary assumption.

Joint work with Riddhi Basu and Jonathan Hermon
Characterization of cutoff for reversible Markov chains
Beyond trees

- The tree assumption can be relaxed. In particular, we can treat jumps to vertices of bounded distance on a tree (i.e. the length of the path from $u$ to $v$ in the tree (which is now just an auxiliary structure) is $> r \implies P(u, v) = 0$) under some mild necessary assumption.

- Previously the BD assumption could not be relaxed mainly due to it being exploited via a representation of hitting times result for BD chains.
The tree assumption can be relaxed. In particular, we can treat jumps to vertices of bounded distance on a tree (i.e. the length of the path from \( u \) to \( v \) in the tree (which is now just an auxiliary structure) is \( > r \implies P(u, v) = 0 \)) under some mild necessary assumption.

Previously the BD assumption could not be relaxed mainly due to it being exploited via a representation of hitting times result for BD chains.

In particular, if \( P(u, v) \geq \delta > 0 \) for all \( u, v \) s.t. \( d_T(u, v) \leq r \) (and otherwise \( P(u, v) = 0 \)), then
\[ \text{cutoff} \iff \text{the Prod. Cond. holds.} \]
Last remark:

- Previously “good expansion of small sets can improve mixing”.

- Now know - considering expansion only of small sets and $t_{rel}$ suffices to bound $t_{mix}$!

\[
    t_{mix}(\epsilon) \leq \text{hit}_{1-\epsilon/4}(3\epsilon/4) + \frac{3t_{rel}}{2} \log \left(\frac{4}{\epsilon}\right).
\]

From which it follows that

\[
    t_{mix} \leq 5 \max_{x,A: \pi(A) \geq 1 - \epsilon/4} \mathbb{E}_x[T_A] + \frac{3t_{rel}}{2} \log \left(\frac{4}{\epsilon}\right).
\]

- For any $x$ and $A$ with $\pi(A) \geq 1 - \epsilon/4$ we can bound $\mathbb{E}_x[T_A]$ using the expansion profile of sets only of $\pi$ measure at most $\epsilon/4$ (by an integral of the form used to bound the mixing time via the expansion profile).
Previously “good expansion of small sets can improve mixing”.

Now know - considering expansion only of small sets and $t_{rel}$ suffices to bound $t_{mix}$!

$$t_{mix}(\epsilon) \leq \text{hit}_{1-\epsilon/4}(3\epsilon/4) + \frac{3t_{rel}}{2} \log \left(\frac{4}{\epsilon}\right).$$

From which it follows that

$$t_{mix} \leq 5 \max_{x,A: \pi(A) \geq 1-\epsilon/4} \mathbb{E}_x[T_A] + \frac{3t_{rel}}{2} \log \left(\frac{4}{\epsilon}\right).$$

For any $x$ and $A$ with $\pi(A) \geq 1 - \epsilon/4$ we can bound $\mathbb{E}_x[T_A]$ using the expansion profile of sets only of $\pi$ measure at most $\epsilon/4$ (by an integral of the form used to bound the mixing time via the expansion profile).

In practice, we can take $\epsilon = \exp[-ct_{mix}/t_{rel}]$ to determine $t_{mix}$ up to a constant.
What can be said about the geometry of the “worst” sets in some interesting particular cases (say, transitivity or monotonicity)?

Joint work with Riddhi Basu and Jonathan Hermon
Open problems

- What can be said about the geometry of the “worst” sets in some interesting particular cases (say, transitivity or monotonicity)?

- When can the worst sets be described as $\{|f_2| \leq C\}$ ($Pf_2 = \lambda_2 f_2$)? (would imply several new cutoff results if true in certain cases)

- When can one relate escaping time from balls of $\pi$-measure $\epsilon$ to escaping time from sets of $\pi$-measure $\epsilon^{100}/100$?

- When can monotonicity w.r.t. a partial order (preserved by the chain) be used to describe the “worst” sets and their hitting time distributions?

Joint work with Riddhi Basu and Jonathan Hermon