Characterization of cutoff for reversible Markov chains

Yuval Peres

Joint work with Riddhi Basu and Jonathan Hermon

3 December 2014

Notation

- Transition matrix - P (reversible).
- Stationary dist. - π.
- Reversibility: $\pi(x)P(x,y) = \pi(y)P(y,x)$, $\forall x, y \in \Omega$.
- Laziness $P(x,x) \geq 1/2$, $\forall x \in \Omega$.

TV distance

- For any 2 dist. μ, ν on Ω, their ***total-variation distance*** is:

$$\|\mu - \nu\|_{\mathrm{TV}} \stackrel{d}{=} \max_{A \subset \Omega} \mu(A) - \nu(A) \ .$$

-

$$d(t,x) \stackrel{d}{=} \|\mathrm{P}_x^t - \pi\|_{\mathrm{TV}}, \qquad d(t) \stackrel{d}{=} \max_{x \in \Omega} d(t,x).$$

TV distance

- For any 2 dist. μ, ν on Ω, their ***total-variation distance*** is:

$$\|\mu - \nu\|_{\mathrm{TV}} \overset{d}{=} \max_{A \subset \Omega} \mu(A) - \nu(A) \; .$$

- $$d(t,x) \overset{d}{=} \|\mathrm{P}_x^t - \pi\|_{\mathrm{TV}}, \qquad d(t) \overset{d}{=} \max_{x \in \Omega} d(t,x).$$

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

$$t_{\mathrm{mix}}(\epsilon) \overset{d}{=} \min\{t : d(t) \le \epsilon\}$$

- $$t_{\mathrm{mix}} \overset{d}{=} t_{\mathrm{mix}}(1/4).$$

Cutoff - definition

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

$$t_{\mathrm{mix}}^{(n)}(\epsilon) - t_{\mathrm{mix}}^{(n)}(1-\epsilon) = o(t_{\mathrm{mix}}^{(n)}), \ \forall\, 0 < \epsilon < 1/4. \qquad (1)$$

Cutoff - definition

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

$$t_{\mathrm{mix}}^{(n)}(\epsilon) - t_{\mathrm{mix}}^{(n)}(1-\epsilon) = o(t_{\mathrm{mix}}^{(n)}), \ \forall\, 0 < \epsilon < 1/4. \tag{1}$$

- (w_n) is called a ***cutoff window*** for $(X_t^{(n)})$ if: $w_n = o\left(t_{\mathrm{mix}}^{(n)}\right)$, and

$$t_{\mathrm{mix}}^{(n)}(\epsilon) - t_{\mathrm{mix}}^{(n)}(1-\epsilon) \le c_\epsilon w_n, \ \forall n \ge 1, \forall \epsilon \in (0, 1/4).$$

Cutoff

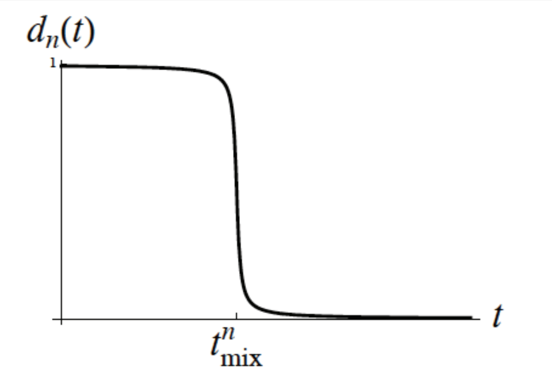


Figure : cutoff

Background

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

Background

- 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.

Background

- 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.

Background

- 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.

Spectral gap & relaxation-time

- Let λ_2 be the largest non-trivial e.v. of P.
- Definition: $\mathrm{gap} = 1 - \lambda_2$ - the **spectral gap**.
- Def: $t_{\mathrm{rel}} := \mathrm{gap}^{-1}$ - the ***relaxation-time***.

The product condition (Prod. cond.)

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

The product condition (Prod. cond.)

- In a 2004 Aim workshop I proposed that ***The product condition (Prod. Cond.)*** - $\mathrm{gap}^{(n)} t_{\mathrm{mix}}^{(n)} \to \infty$ (equivalently, $t_{\mathrm{rel}}^{(n)} = o(t_{\mathrm{mix}}^{(n)})$) 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. **Prod. Cond.** $\Longrightarrow$ cutoff.

Aldous' example

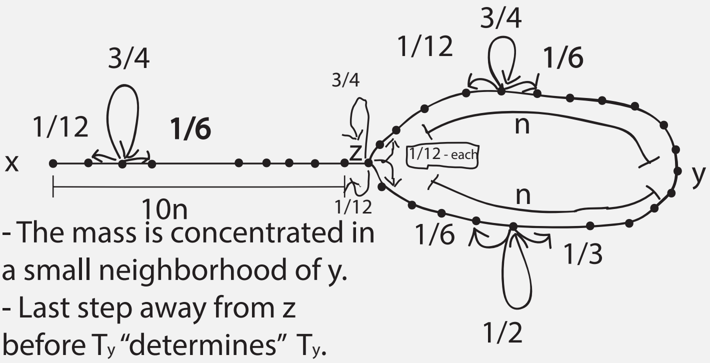
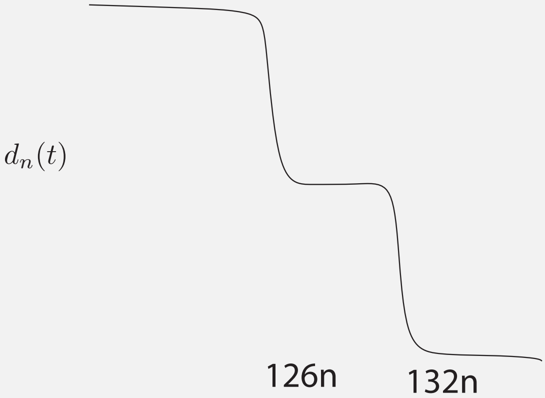


Figure : Fixed bias to the right conditioned on a non-lazy step.

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

Aldous' example



Hitting and Mixing

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

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 $0 < \alpha \leq 1/2$:

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

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 $0 < \alpha \leq 1/2$:

$$t_{\mathrm{H}}(\alpha) = \Theta_\alpha(t_{\mathrm{mix}}), \text{ where } t_{\mathrm{H}}(\alpha) := \max_{x,A:\, \pi(A) \geq \alpha} \mathbb{E}_x[T_A]. \qquad (2)$$

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

counter-example

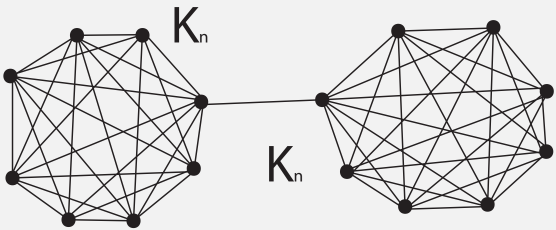


Figure : n is the index of the chain

Hitting and Cutoff

- Concentration of hitting times of “worst” sets is related to cutoff in birth and death (**BD**) chains.

Hitting and Cutoff

- 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.

Hitting and Cutoff

- 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, π) *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, π) *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* $c_\epsilon = C\sqrt{|\log \epsilon|}$.

- 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 $w_n = \Omega\left(\sqrt{t_{\text{rel}}^{(n)} t_{\text{mix}}^{(n)}}\right)$.

To mix - escape and then relax

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

$$\mathrm{hit}_{\alpha,x}(\epsilon) := \min\{t : \mathrm{P}_x[T_A > t] \le \epsilon : \text{for all } A \subset \Omega \text{ s.t. } \pi(A) \ge \alpha\},$$

$$\mathrm{hit}_{\alpha}(\epsilon) := \max_{x \in \Omega} \mathrm{hit}_{\alpha,x}(\epsilon)$$

To mix - escape and then relax

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

$$\mathrm{hit}_{\alpha,x}(\epsilon) := \min\{t : \mathrm{P}_x[T_A > t] \le \epsilon : \text{for all } A \subset \Omega \text{ s.t. } \pi(A) \ge \alpha\},$$

$$\mathrm{hit}_\alpha(\epsilon) := \max_{x\in\Omega} \mathrm{hit}_{\alpha,x}(\epsilon)$$

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

To mix - escape and then relax

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

$$\mathrm{hit}_{\alpha,x}(\epsilon) := \min\{t : \mathrm{P}_x[T_A > t] \le \epsilon : \text{for all } A \subset \Omega \text{ s.t. } \pi(A) \ge \alpha\},$$

$$\mathrm{hit}_\alpha(\epsilon) := \max_{x \in \Omega} \mathrm{hit}_{\alpha,x}(\epsilon)$$

- 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).

Hitting times when $X_0 \sim \pi$

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

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

Hitting times when $X_0 \sim \pi$

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

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

- By a coupling argument,

$$\mathrm{P}_x[T_A > t + 2st_{\mathrm{rel}}] \leq d(t) + \mathrm{P}_\pi[T_A > 2st_{\mathrm{rel}}].$$

Hitting of worst sets

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

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

- Terms involving t_{rel} are negligible under the Prod. Cond..

Hitting of worst sets

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

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

- Terms involving t_{rel} are negligible under the Prod. Cond..

Hitting of worst sets

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

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

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

Main abstract result

Definition: A sequence has hit_α-cutoff if

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

Main abstract result

Definition: A sequence has hit_α-cutoff if

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

Main abstract result:

Theorem

Let (Ω_n, P_n, π_n) be a seq. of finite reversible lazy MCs. Then TFAE:

- *The seq. exhibits cutoff.*
- *The seq. exhibits a hit_α-cutoff for some $\alpha \in (0, 1/2)$.*
- *The seq. exhibits a hit_α-cutoff for some $\alpha \in [1/2, 1)$ and the Prod. Cond. holds.*

Main abstract result

Definition: A sequence has hit_α-cutoff if

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

Main abstract result:

Theorem

Let (Ω_n, P_n, π_n) be a seq. of finite reversible lazy MCs. Then TFAE:

- *The seq. exhibits cutoff.*
- *The seq. exhibits a hit_α-cutoff for some $\alpha \in (0, 1/2)$.*
- *The seq. exhibits a hit_α-cutoff for some $\alpha \in [1/2, 1)$ and the Prod. Cond. holds.*

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

Main abstract result

Definition: A sequence has hit_α-cutoff if

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

Main abstract result:

Theorem

Let (Ω_n, P_n, π_n) be a seq. of finite reversible lazy MCs. Then TFAE:

- *The seq. exhibits cutoff.*
- *The seq. exhibits a hit_α-cutoff for some $\alpha \in (0, 1/2)$.*
- *The seq. exhibits a hit_α-cutoff for some $\alpha \in [1/2, 1)$ and the Prod. Cond. holds.*

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

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

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

Tools

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

Tools

- 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 $\|f\|_2^2 := \mathbb{E}_{\pi}[f^2]$.

Tools

- 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 $\mathrm{Var}_\pi g := \|g - \mathbb{E}_\pi[g]\|_2^2$.

Tools

- 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 $\mathrm{Var}_\pi g := \|g - \mathbb{E}_\pi[g]\|_2^2$.

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

Lemma (Contraction Lemma)

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

$$\mathrm{Var}_\pi P^t 1_A \leq e^{-2t/t_{\mathrm{rel}}}. \quad (3)$$

Maximal Inequality

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

Theorem (Maximal inequality)

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

$$f^*(x) := \sup_{0 \le k < \infty} |P^k(f)(x)| = \sup_{0 \le k < \infty} |E_x[f(X_k)]|.$$

Then for $1 < p < \infty$*,*

$$\|f^*\|_p \le q\|f\|_p \qquad 1/p + 1/q = 1 \tag{4}$$

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 ϵ-mixed w.r.t. A" and to control its π 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. $T_G \le t$ serves as a certificate of "being ϵ-mixed w.r.t. A" and to control its π measure from below.

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

$$G = G_s(A) := \left\{ g : \forall \tilde{s} \ge s, |\mathrm{P}_g^{\tilde{s}}(A) - \pi(A)| \le 4\sigma_s \right\} .$$

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 ϵ-mixed w.r.t. A" and to control its π measure from below.

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

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. $T_G \le t$ serves as a certificate of "being ϵ-mixed w.r.t. A" and to control its π measure from below.

- Let $\sigma_s := e^{-s/t_{\text{rel}}} \ge \sqrt{\text{Var}_\pi P^s 1_A}$ (contraction lemma).
- Consider
$$G = G_s(A) := \left\{ g : \forall \tilde{s} \ge s, |\mathrm{P}_g^{\tilde{s}}(A) - \pi(A)| \le 4\sigma_s \right\} .$$
- Want precision $4\sigma_s = \epsilon \implies s := t_{\text{rel}} \times \log(4/\epsilon)$.

Claim

$$\pi(G) \ge 1/2. \quad (5)$$

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 ϵ-mixed w.r.t. A" and to control its π measure from below.

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

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

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

Claim

$$\pi(G) \geq 1/2. \tag{5}$$

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

$$G^c \subset \{f_s^* > 4\|f_s\|_2\} .$$

Apply Starr's inequality.

Main idea

- Claim:
$$t_{\mathrm{mix}}(2\epsilon) \le \mathrm{hit}_{1/2}(\epsilon) + t_{\mathrm{rel}} \times \log(4/\epsilon).$$
- *Proof:* Recall
$$G := G_s(A,m) := \left\{ g : \forall \tilde{s} \ge s, |\mathrm{P}_g^{\tilde{s}}(A) - \pi(A)| \le \epsilon \right\}, \; s := t_{\mathrm{rel}} \times \log(4/\epsilon)$$
- Set $t := \mathrm{hit}_{1/2}(\epsilon)$. By prev. claim $\pi(G) \ge 1/2 \Longrightarrow \mathrm{P}_x[T_G > t] \le \epsilon$ (by def. of t).

Main idea

- Claim:
$$t_{\mathrm{mix}}(2\epsilon) \le \mathrm{hit}_{1/2}(\epsilon) + t_{\mathrm{rel}} \times \log(4/\epsilon).$$
- *Proof:* Recall
$$G := G_s(A,m) := \left\{ g : \forall \tilde{s} \ge s, |\mathrm{P}_g^{\tilde{s}}(A) - \pi(A)| \le \epsilon \right\}, \; s := t_{\mathrm{rel}} \times \log(4/\epsilon)$$
- Set $t := \mathrm{hit}_{1/2}(\epsilon)$. By prev. claim $\pi(G) \ge 1/2 \Longrightarrow \mathrm{P}_x[T_G > t] \le \epsilon$ (by def. of t).
- For any x, A:
$$|\mathrm{P}_x^{t+s}(A) - \pi(A)| \le \mathrm{P}_x[T_G > t] + \max_{g \in G, \tilde{s} \ge s} |\mathrm{P}_g^{\tilde{s}}(A) - \pi(A)| \le 2\epsilon.$$

Trees

- Let: $T := (V, E)$ be a finite tree.
- (V, P, π) 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.

Trees

- Can the tree structure be used to determine the identity of the “worst” sets?

Trees

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

Trees

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

Trees

How to generalize this to trees?

Central vertex

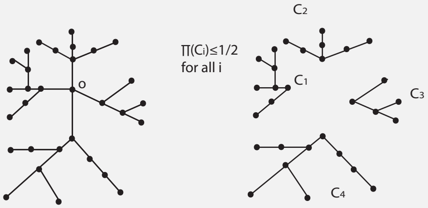


Figure : A vertex $o \in V$ is called a ***central-vertex*** if each connected component of $T \setminus \{o\}$ has stationary probability at most 1/2.

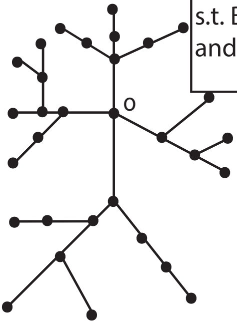
Trees

- There is always a central-vertex (and at most 2). We fix one, denote it by o and call it the **root**.

Trees

- 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 $\Longrightarrow \exists$ such unweighed trees (Peres-Sousi (13)).

Trees



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 Π_B is Π
conditioned on B.

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

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

Trees

- Cutoff would follow if we show that T_o is concentrated (under the Prod. Cond.).
- More precisely, we need to show that $\mathbb{E}_x[T_o] = \Omega(t_{\mathrm{mix}}) \Longrightarrow T_{y_\beta(x)}$ is concentrated if $X_0 = x$.

Trees

- Cutoff would follow if we show that T_o is concentrated (under the Prod. Cond.).
- More precisely, we need to show that $\mathbb{E}_x[T_o] = \Omega(t_{\mathrm{mix}}) \Longrightarrow T_{y_\beta(x)}$ is concentrated if $X_0 = x$.

Trees

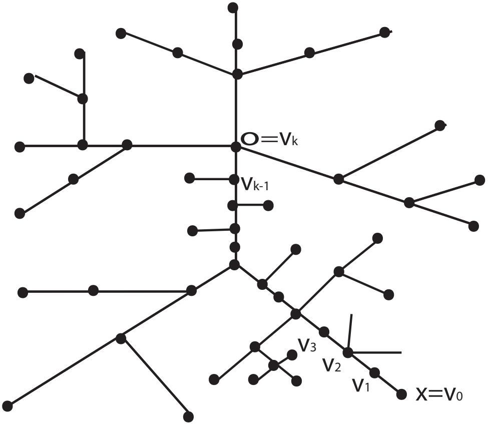


Figure : Let $v_0 = x, v_1, \ldots, v_k = o$ be the vertices along the path from x to o.

Trees

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

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

Trees

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

- It suffices to show that $\mathrm{Var}_x[T_o] \leq 4t_{\mathrm{rel}}\mathbb{E}_x[T_o]$.
- If $X_0 = x$ then T_o is the sum of crossing times of the edges along the path between x: $\tau_i := T_{v_i} - T_{v_{i-1}} \stackrel{d}{=} T_{v_i}$ under $X_0 = v_{i-1}$

Trees

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

- It suffices to show that $\mathrm{Var}_x[T_o] \le 4 t_{\mathrm{rel}} \mathbb{E}_x[T_o]$.
- If $X_0 = x$ then T_o is the sum of crossing times of the edges along the path between x: $\tau_i := T_{v_i} - T_{v_{i-1}} \stackrel{d}{=} T_{v_i}$ under $X_0 = v_{i-1}$
- $\tau_1, \dots, \tau_k$ are independent $\implies$ it suffices to bound the sum of their 2nd moments $\mathrm{Var}_x[T_o] = \sum \mathrm{Var}_x[\tau_i] = \sum \mathrm{Var}_{v_{i-1}}[T_{v_i}] \le \sum \mathbb{E}_{v_{i-1}}[T_{v_i}^2]$.

Trees

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

- It suffices to show that $\mathrm{Var}_x[T_o] \le 4 t_{\mathrm{rel}} \mathbb{E}_x[T_o]$.
- 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$X_0 = v_{i-1}$
- $\tau_1, \dots, \tau_k$ are independent $\Longrightarrow$ it suffices to bound the sum of their 2nd moments $\mathrm{Var}_x[T_o] = \sum \mathrm{Var}_x[\tau_i] = \sum \mathrm{Var}_{v_{i-1}}[T_{v_i}] \le \sum \mathbb{E}_{v_{i-1}}[T_{v_i}^2]$.
- 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 π_A be π conditioned on A.
- Kac formula implies that for any A, there exists a dist. μ on the external vertex boundary of A s.t. $\mathrm{E}_\mu[T_A^2] \le 2\mathrm{E}_\mu[T_A]\mathrm{E}_{\pi_{A^c}}[T_A] \Longrightarrow$
- By the tree structure $\mathrm{E}_{v_{i-1}}[T_{v_i}^2] \le 2\mathrm{E}_{v_{i-1}}[T_{v_i}]\mathrm{E}_{\pi_{W_{v_{i-1}}}}[T_{v_i}]$.
- Not hard to show $\mathrm{E}_{\pi_{W_{v_{i-1}}}}[T_{v_i}] \le 2t_{\mathrm{rel}}$ (generally $\pi(A^c)\mathrm{E}_{\pi_{A^c}}[T_A] \le t_{\mathrm{rel}}$) so

$$\sum \mathrm{E}_{v_{i-1}}[T_{v_i}^2] \le \sum 4t_{\mathrm{rel}}\mathrm{E}_{v_{i-1}}[T_{v_i}] = 4t_{\mathrm{rel}}\mathbb{E}_x[T_o]. \quad \square$$

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.

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.

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.

- 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}!

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

From which it follows that

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

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

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}!

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

  From which it follows that

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

- 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 π 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 $\epsilon = \exp[-c t_{\mathrm{mix}}/t_{\mathrm{rel}}]$ to determine t_{mix} up to a constant.

Open problems

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

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 π-measure ϵ to escaping time from sets of π-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?