# Restart perturbations for reversible Markov chains

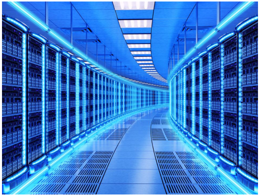
## Trichotomy and pre-cutoff equivalence

Daniel Vial, Vijay Subramanian

ECE Department, University of Michigan

# Motivation (1/2)

Discrete-time Markov chain (DTMC): a common stochastic model



Fundamental concern: sensitivity of stationary distribution to (bounded) perturbations of transition matrix

- e.g. how much do modeling errors affect long-run behavior?
- e.g. how much damage can adversary do long-term?

# Motivation (2/2)

Sensitivity and mixing time

- Mixing time: number steps for DTMC to reach equilibrium/stationarity
- If fast mixing, less time for perturbation to take effect
- If slow mixing, small perturbation may have large effect

In particular, connection to cutoff

- Cutoff: far from stationary for many steps, then suddenly stationary

**Table 1. Distance to stationarity for repeated shuffles of 52 cards**

| | $k$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $\|P^k - \pi\|$ | 1.000 | 1.000 | 1.000 | 1.000 | 0.924 | 0.624 | 0.312 | 0.161 | 0.083 | 0.041 |

(from Diaconis 1996)

- Many examples, but little general theory

# Overview

1. Perturbation model
2. Mixing times and cutoff
3. Illustrative examples
4. Results (and related work)

# Basic definitions

DTMC $\{X_n(t)\}_{t\in\mathbb{Z}_+}$ with states $[n] = \{1, \ldots, n\}$ and transition matrix $P_n$:

$$\mathbb{P}(X_n(t+1) = j | X_n(t) = i) = P_n(i,j) \;\forall\; i,j \in [n], t \in \mathbb{Z}_+$$

Assume irreducible, aperiodic $\Rightarrow$ converge to unique stationary distribution $\pi_n$ from any starting state:

$$\lim_{t\to\infty} P_n^t(i,\cdot) = \pi_n = \pi_n P_n \;\forall\; i \in [n]$$

Sometimes require reversibility (a.k.a. local balance):

$$\pi_n(i)P_n(i,j) = \pi_n(j)P_n(j,i) \;\forall\; i,j \in [n]$$

Sometimes require laziness (but believe we can drop this):

$$P_n(i,i) \geq \frac{1}{2} \;\forall\; i \in [n]$$

# Perturbation model

Given $\alpha_n \in (0,1)$, define

$$B(P_n, \alpha_n) = \left\{ \tilde{P}_n \in \mathcal{E}_n : \max_{i \in [n]} \|\tilde{P}_n(i,\cdot) - P_n(i,\cdot)\|_{TV} \leq \alpha_n \right\}$$

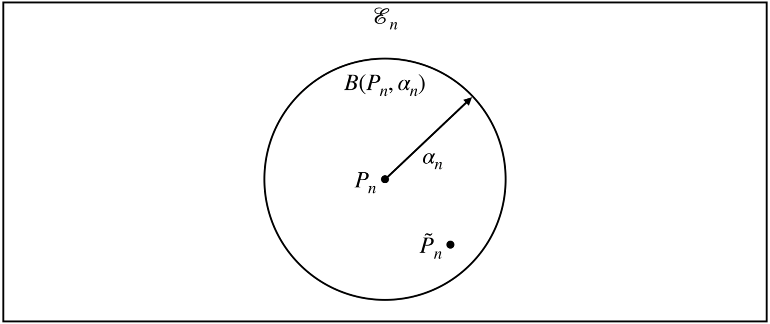
where

- $\mathcal{E}_n$ = set of $n \times n$ irreducible/aperiodic transition matrices
- $\|\mu - \nu\|_{TV} = \max_{A \subset [n]} |\mu(A) - \nu(A)|$ for distributions $\mu, \nu$ on $[n]$

Denote stationary distribution of $\tilde{P}_n \in \mathcal{E}_n$ by $\tilde{\pi}_n$

**Main focus of this work:** relationship between $\|\tilde{\pi}_n - \pi_n\|_{TV}$ and $\alpha_n$
i.e. how does steady-state error scale with perturbation magnitude?

# Illustration and interpretations




| | **Interpretation 1** | **Interpretation 2** |
|---|---|---|
| $P_n$ | True dynamics | Designed system |
| $\tilde{P}_n$ | Modeled dynamics | System disrupted by adversary |
| $\alpha_n$ | Modeling accuracy | Adversary budget/ability |
| $\lVert\tilde{\pi}_n - \pi_n\rVert_{TV}$ | Effect of modeling errors | Effect of adversary |

# Restart perturbations

Given $\alpha_n \in (0,1)$ and distribution $\sigma_n$ over $[n]$, define

$$P_{\alpha_n,\sigma_n} = (1-\alpha_n)P_n + \alpha_n 1_n^{\mathsf{T}} \sigma_n \in B(P_n, \alpha_n)$$

and denote corresponding stationary distribution by $\pi_{\alpha_n,\sigma_n}$

Interpretation:

- Flip Bernoulli($\alpha_n$) coin at each step
- If tails, sample next state from $P_n$
- If heads, sample next state from $\sigma_n$

Note: $\pi_{\alpha_n,\pi_n} = \pi_n$ (perturbations need not change stationary distribution)

# Properties of restart perturbations

1. Power iteration: Solving $\pi_{\alpha_n,\sigma_n} = \pi_{\alpha_n,\sigma_n} P_{\alpha_n,\sigma_n}$ gives

$$\pi_{\alpha_n,\sigma_n} = \alpha_n \sigma_n (I_n - (1-\alpha_n) P_n)^{-1} = \alpha_n \sigma_n \sum_{t=0}^{\infty} (1-\alpha_n)^t P_n^t$$

2. Linearity: For distributions $\sigma_{n,1}, \sigma_{n,2}$ and $\lambda \in (0,1)$,

$$\pi_{\alpha_n, \lambda\sigma_{n,1} + (1-\lambda)\sigma_{n,2}} = \lambda \pi_{\alpha_n,\sigma_{n,1}} + (1-\lambda)\pi_{\alpha_n,\sigma_{n,2}}$$

3. Perfect sampling: Letting $T \sim \text{Geometric}(\alpha_n)$,

$$\pi_{\alpha_n,\sigma_n}(i) = \mathbb{P}(X_n(T) = i | X_n(0) \sim \sigma_n)$$

# Sidebar: (Personalized) PageRank

When $P_n$ is random walk on $G = ([n], E)$ . . .

- $\pi_{\alpha_n, 1_n/n}$ called PageRank, measure of importance/centrality of nodes
- $\pi_{\alpha_n, e_i}$ called Personalized PageRank for $i$, measure of relevance to $i$

Applications: Internet search (Page et al. 1999), recommendation systems (Gupta et al. 2013), community detection (Andersen, Chung, Lang 2006; Kloumann, Ugander, Kleinberg 2017), graph similarity (Koutra, Vogelstein, Faloutsos 2013), bioinformatics (Morrison et al. 2005; Freschi 2007), . . .

Rich literature: efficient PageRank estimation (Jeh, Widom 2003; Avrachenkov et al. 2007; Andersen et al. 2008; Lofgren, Banerjee, Goel 2016; Vial, Subramanian 2017), implementation e.g. MapReduce (Tong, Faloutsos, Pan 2006; Kang, Tsourakakis, Faloutsos 2009; Whang et al. 2015), behavior on random graphs (Avrachenkov et al. 2015; Chen, Litvak, Olvera-Cravioto 2017; Caputo, Quattropani 2019; Vial, Subramanian 2019), . . .

# Mixing times

Define $t$-step distance from stationarity and $\epsilon$-mixing time as

$$d_n(t) = \max_{i \in [n]} \|P_n^t(i, \cdot) - \pi_n\|_{TV}, \quad t_{\text{mix}}^{(n)}(\epsilon) = \min\{t \in \mathbb{Z}_+ : d_n(t) \leq \epsilon\}$$

Also define mixing time $t_{\text{mix}}^{(n)} = t_{\text{mix}}^{(n)}(\frac{1}{4})$

- Well-known fact: $d_n(kt_{\text{mix}}^{(n)}(\epsilon)) \leq (2\epsilon)^k$; $\epsilon = \frac{1}{4}$ gives $d_n(kt_{\text{mix}}^{(n)}) \leq 2^{-k}$

Typically desire "Big-O" behavior of $t_{\text{mix}}^{(n)}$ for sequence $\{P_n\}_{n \in \mathbb{N}}$

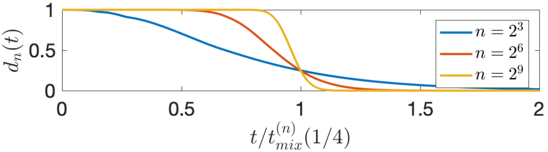
- e.g. $t_{\text{mix}}^{(n)} = \Theta(n^2)$ for random walk on $n$-cycle

More refined question: how "sharply" does graph of $d_n(t)$ decay to 0?

# (Pre-)cutoff

Sequence of chains exhibits cutoff if

$$\lim_{n\to\infty} \frac{t_{\text{mix}}^{(n)}(\epsilon)}{t_{\text{mix}}^{(n)}(1-\epsilon)} = 1 \ \forall \ \epsilon \in \left(0, \frac{1}{2}\right)$$




Weaker condition of pre-cutoff only requires

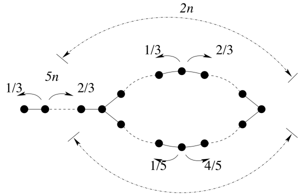
$$\sup_{\epsilon\in(0,\frac{1}{2})} \underbrace{\limsup_{n\to\infty} \frac{t_{\text{mix}}^{(n)}(\epsilon)}{t_{\text{mix}}^{(n)}(1-\epsilon)}}_{\star} < \infty$$

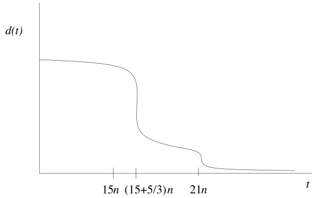
# Literature on cutoff (1/2)

Many examples of chains with cutoff

- Classic examples (Diaconis 1996): card shuffling (e.g. Diaconis, Shahshahani 1981; Bayer, Diaconis 1992), diffusion models (e.g. Aldous 1983; Diaconis, Shahshahani 1987), . . .
- Recently, cutoff for generative models, e.g. random walks on random regular graphs (Lubetzky, Sly 2010), non-backtracking walks on sparse random graphs (Ben-Hamou, Salez 2017), general models of sparse chains (Bordenave, Caputo, Salez 2019), . . .

Example of pre-cutoff but no cutoff (Levin, Peres, Wilmer 2009, Ch. 18):



# Literature on cutoff (2/2)

Necessary condition for pre-cutoff (and thus cutoff): inverse spectral gap vanishes relative to $t_{\text{mix}}^{(n)}$ (Levin, Peres, Wilmer 2009, Prop. 18.4)

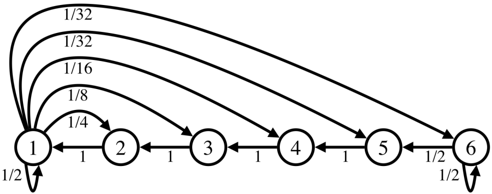
Spectral gap condition not sufficient (Levin, Peres, Wilmer 2009, Ex. 18.7)

No known necessary/sufficient condition in terms of eigenvalues alone (Diaconis 1996 conjectured high multiplicity of second eigenvalue)

Basu, Hermon, Peres 2015: cutoff equivalent to “hitting time cutoff”

**Necessary/sufficient conditions for pre-cutoff?**

# Winning streak reversal (WSR) from Levin, Peres, Wilmer 2009




$$P_n(i,j) = \begin{cases} 2^{-j}, & i = 1, j \in \{1, \dots, n-1\} \\ 2^{-(n-1)}, & i = 1, j = n \\ 1, & i \in \{2, \dots, n-1\}, j = i-1 \\ 2^{-1}, & i = n, j \in \{n-1, n\} \end{cases}$$

$\pi_n = [2^{-1}\ 2^{-2}\ \cdots\ 2^{-(n-1)}\ 2^{-(n-1)}]$, since

$$\sum_{j=1}^{n} \pi_n(j) P_n(j,i) = 2^{-1} \times 2^{-i} + 2^{-(i+1)} \times 1 = 2^{-i} = \pi_n(i)$$

## WSR mixing behavior

Stationary by step $n-1$, i.e. $t_{\text{mix}}^{(n)}(1-\epsilon), t_{\text{mix}}^{(n)}(\epsilon) \le n-1$

- $P_n(1,\cdot) = \pi_n$ by previous slide
- For $i \in \{2,\dots,n-1\}$, move to state 1 in $i-1$ steps, so $P_n^i(i,\cdot) = \pi_n$
- More subtly, $P_n^{n-1}(n,\cdot) = \pi_n$

Lower bound: $t_{\text{mix}}^{(n)}(1-\epsilon), t_{\text{mix}}^{(n)}(\epsilon) \ge n-1+\log_2(\epsilon)$, since

$$\begin{aligned} d_n(t) &= \max_{i\in[n], A\subset[n]} |P_n^t(i,A) - \pi_n(A)| \\ &\ge P_n^t(n-1,\{n-1-t\}) - \pi_n(\{n-1-t\}) = 1-2^{-n+1+t} \end{aligned}$$

Combining gives strong form of cutoff:

$$1 \le \frac{t_{\text{mix}}^{(n)}(\epsilon)}{t_{\text{mix}}^{(n)}(1-\epsilon)} = 1 + O\left(\frac{1}{n}\right)$$

# WSR perturbation behavior

Intuitively, "worst" perturbation is restart at $n$, i.e. $P_{\alpha_n, e_n}$

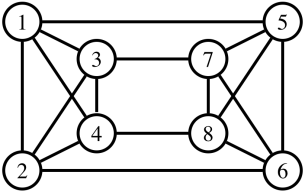
Suppose $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to \infty$ (e.g. $\alpha_n = 1/\sqrt{n}$):

- Restart at $n$ every $\sqrt{n}$ steps $\Rightarrow$ rarely visit $[n/2] \Rightarrow \pi_{\alpha_n, e_n}([n/2]) \approx 0$
- Also $\pi_n([n/2]) = \sum_{i=1}^{n/2} 2^{-i} \approx 1$ (for large $n$)
- Taken together, $\|\pi_n - \pi_{\alpha_n, e_n}\|_{TV} \geq \pi_n([n/2]) - \pi_{\alpha_n, e_n}([n/2]) \approx 1$

If instead $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to 0$ (e.g. $\alpha_n = 1/n^2$):

- Visit state 1 before restart, reach original stationary distribution
- Once "locked into" this distribution, can't escape
- Consequently, $\|\pi_n - \pi_{\alpha_n, e_n}\|_{TV} \approx 0$

# Complete graph bijection (CGB)




Construction:

- Draw complete graphs / cliques on $\{1, \ldots, \frac{n}{2}\}$ and $\{1 + \frac{n}{2}, \ldots, n\}$
- Add edges $\{(1, 1 + \frac{n}{2}), (2, 2 + \frac{n}{2}), \ldots, (\frac{n}{2}, n)\}$

We study lazy simple random walk on resulting graph

By symmetry, uniform distribution is stationary

## CGB mixing behavior

"Half-mixed" in one step, i.e. $t_{\text{mix}}^{(n)}(1-\epsilon) = \Theta(1)$

- $P_n(X_n(0), \cdot) \approx$ uniform on $X_n(0)$'s clique $\Rightarrow \|P_n(X_n(0), \cdot) - \pi_n\|_{TV} \approx \frac{1}{2}$

To fully mix, require $n$ steps, i.e. $t_{\text{mix}}^{(n)}(\epsilon) = \Theta(n)$

- At each step, $\Theta(\frac{1}{n})$ mass escapes $X_n(0)$'s clique
- Thus, $\Theta(n)$ steps for constant mass to escape $X_n(0)$'s clique

"Opposite" of cutoff, since

$$\frac{t_{\text{mix}}^{(n)}(\epsilon)}{t_{\text{mix}}^{(n)}(1-\epsilon)} = \Theta(n)$$

which is maximal (up to constants) among $t_{\text{mix}}^{(n)}(\epsilon) = \Theta(n)$ chains

# CGB perturbation behavior

By symmetry, "worst" (restart) perturbation is $P_{\alpha_n, e_1}$

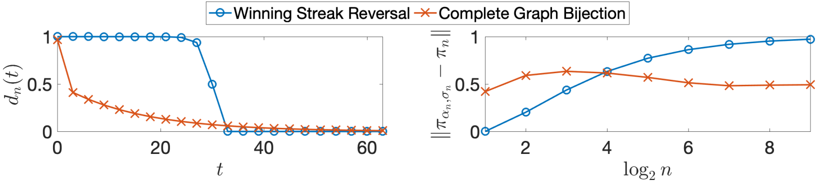
One step after restart, uniform on $\{1, \ldots, \frac{n}{2}\}$

"Half-mixed", can't fully escape this distribution

More specifically, can show that for any $\alpha_n \to 0$ and any $\tilde{P}_n \in B(P_n, \alpha_n)$,

$$\limsup_{n \to \infty} \|\tilde{\pi}_n - \pi_n\|_{TV} \leq \frac{1}{2}$$

# Summary and key insights




WSR: strong form of cutoff, sensitive to $\alpha_n$-magnitude perturbations

- Assuming $\alpha_n \gg 1/t_{\text{mix}}^{(n)}(\epsilon)$

CGB: "opposite" of cutoff, robust against perturbation

Do these phenomena hold more generally?

# Trichotomy result

## Theorem (Vial, Subramanian 2019)

*Let* $P_n \in \mathcal{E}_n, \alpha_n \in (0,1), \epsilon \in (0,1)$. *Assume* $\lim_{n\to\infty} \alpha_n t_{mix}^{(n)}(\epsilon) = c \in [0,\infty]$, $P_n$ *is lazy and reversible* $\forall$ $n$, *and* $\{P_n\}$ *has cutoff.*

- *If* $c = 0$, *then* $\forall$ $\{\tilde{P}_n\}$ *with* $\tilde{P}_n \in B(P_n, \alpha_n)$ $\forall$ $n$,

$$\lim_{n\to\infty} \|\tilde{\pi}_n - \pi_n\|_{TV} = 0.$$

- *If* $c = \infty$, *then* $\exists$ $\{\sigma_n\}$ *with* $\sigma_n \in \Delta_{n-1}$ $\forall$ $n$ *s.t.*

$$\lim_{n\to\infty} \|\pi_{\alpha_n,\sigma_n} - \pi_n\|_{TV} = 1.$$

- *If* $c \in (0,\infty)$, *then* $\forall$ $\{\sigma_n\}$ *with* $\sigma_n \in \Delta_{n-1}$ $\forall$ $n$,

$$\limsup_{n\to\infty} \|\pi_{\alpha_n,\sigma_n} - \pi_n\|_{TV} \leq 1 - e^{-c}.$$

  *Also, the bound is tight, i.e.* $\exists$ $\{\sigma_n\}$ *with* $\sigma_n \in \Delta_{n-1}$ $\forall$ $n$ *s.t.*

$$\liminf_{n\to\infty} \|\pi_{\alpha_n,\sigma_n} - \pi_n\|_{TV} \geq 1 - e^{-c}.$$

# Notes on theorem

Tight characterization: given any $c' \in [0,1]$, can choose $\alpha_n, \sigma_n$ s.t.

$$\lim_{n\to\infty} \|\pi_{\alpha_n,\sigma_n} - \pi_n\|_{TV} = c'.$$

Some bounds require weaker assumptions (see Lemmas 1-2 in paper)

- Upper bounds don't need laziness, reversibility, or cutoff
- Lower bounds only need pre-cutoff in the case $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to \infty$

Notes on assumptions for lower bound:

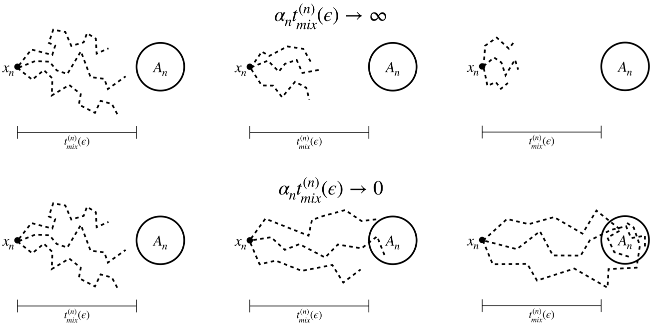
- Believe we can drop laziness
- Reversibility common in mixing times literature (real eigenvalues)
- Necessity of (pre-)cutoff – next theorem

# Proof idea

Basu, Hermon, Peres 2015: if pre-cutoff holds, $\exists\, x_n \in [n], A_n \subset [n]$ s.t.

$$\pi_n(A_n) \approx 1, \quad t_{\text{mix}}^{(n)}(\epsilon) \propto \text{time to reach } A_n \text{ from } x_n$$

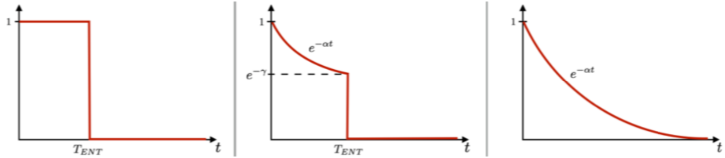
i.e. "mixing times are hitting times of large sets" (also Peres, Sousi 2015)



# Related results in literature (1/2)

Caputo, Quattropani 2019 shows distance to stationarity $d_{\alpha_n,\sigma_n}(t)$ on $P_{\alpha_n,\sigma_n}$ follows trichotomy when $P_n$ = random walk on configuration model (CM):

- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to 0$, $d_{\alpha_n,\sigma_n}$ is step function, like $d_n$ (left)
- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to \infty$, $d_{\alpha_n,\sigma_n}(t)$ decays exponentially (right)
- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to (0,\infty)$, interpolation (middle)




Avena et al. 2018: similar result for dynamic CM

# Related results in literature (2/2)

Vial, Subramanian 2019 studies matrix $\Pi = \{\pi_{\alpha_n, e_i}\}_{i \in [n]}$ for CM

- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to 0$, rank$(\Pi) = \Theta(1)$
- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to \infty$, rank$(\Pi) = \Omega(\frac{n}{\log n})$ (conjecture)
- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to (0, \infty)$, rank$(\Pi) = O(n^c), c \in (0, 1)$

These four results have similar flavor:

- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to 0$, some property unchanged
- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to \infty$, this property maximally changed
- If $\alpha_n t_{\text{mix}}^{(n)}(\epsilon) \to (0, \infty)$, intermediate behavior

Suggests (ill-posed) question: characterizing this class of properties

# Pre-cutoff "equivalence"

Some notion of cutoff assumed for all of our lower bounds – is it necessary?

## Sensitivity Condition (SC)

*For certain sequences* $\{\alpha_{n,\epsilon}\}_{n\in\mathbb{N},\epsilon\in(0,1/2)} \subset (0,1)$, $\exists$ $\{\sigma_{n,\epsilon}\}_{n\in\mathbb{N},\epsilon\in(0,1/2)}$ *such that* $\lim_{n\to\infty} \|\pi_n - \pi_{\alpha_{n,\epsilon},\sigma_{n,\epsilon}}\|_{TV} = 1\ \forall\ \epsilon \in (0,1/2)$.

## Theorem (Vial, Subramanian 2019)

*Assume each* $P_n$ *is lazy and reversible* $\forall\ n$*. If pre-cutoff holds, then SC holds. If*

$$\sup_{\epsilon\in(0,1/2)} \liminf_{n\to\infty} t_{mix}^{(n)}(\epsilon)/t_{mix}^{(n)}(1-\epsilon) = \infty, \qquad (1)$$

*then SC fails.*

- (1) slightly stronger than converse of pre-cutoff (lim inf instead of lim sup)
- Thus, theorem says pre-cutoff and SC (almost) equivalent
- Complements Basu, Hermon, Peres 2015 (cutoff $\Leftrightarrow$ "hitting time cutoff")

# Illustration of "equivalence"

Parition of lazy/reversible sequences induced by SC:

| SC holds | SC fails |
| --- | --- |
| $\sup_{\epsilon\in(0,1/2)} \limsup_{n\to\infty} \frac{t_{mix}^{(n)}(\epsilon)}{t_{mix}^{(n)}(1-\epsilon)} < \infty$ | $\sup_{\epsilon\in(0,1/2)} \liminf_{n\to\infty} \frac{t_{mix}^{(n)}(\epsilon)}{t_{mix}^{(n)}(1-\epsilon)} = \infty$ |

Believe gray area only contains "bizarre" sequences

- e.g. alternating between WSR and CGB

# "Certain sequences"

## Sensitivity Condition (SC)

*For any* $\{\alpha_{n,\epsilon}\}_{n\in\mathbb{N},\epsilon\in(0,1/2)} \subset (0,1)$ *satisfying* (2), $\exists$ $\{\sigma_{n,\epsilon}\}_{n\in\mathbb{N},\epsilon\in(0,1/2)}$ *such that* $\lim_{n\to\infty} \|\pi_n - \pi_{\alpha_{n,\epsilon},\sigma_{n,\epsilon}}\|_{TV} = 1 \ \forall\ \epsilon \in (0,1/2)$.

$$\sup_{\epsilon\in(0,1/2)} \liminf_{n\to\infty} \alpha_{n,\epsilon} t_{mix}^{(n)}(\epsilon) = \infty, \ \frac{\alpha_{n,\epsilon}}{\alpha_{n,\delta}} \in \left[\frac{t_{mix}^{(n)}(1-\delta)}{t_{mix}^{(n)}(1-\epsilon)}, 1\right] \ \forall\ 0 < \delta \le \epsilon < \frac{1}{2}. \quad (2)$$

Proof intuition broadly similar to previous theorem

Little intuition for (2), but yields the following "magic":

- When pre-cutoff fails, $\{\{\alpha_{n,\epsilon}\}_{n,\epsilon} : (2) \text{ holds}\}$ is sufficiently large that we can find $\{\sigma_{n,\epsilon}\}_{n,\epsilon}$ violating conclusion of SC
- When pre-cutoff holds, $\{\{\alpha_{n,\epsilon}\}_{n,\epsilon} : (2) \text{ holds}\}$ collapses to a set for which conclusion of SC holds

# Conclusions

Main takeaways

- Perturbation error depends on perturbation magnitude $\times$ mixing time
- Pre-cutoff almost equivalent to certain notion of "perturbation sensitivity"

Further reading


- Vial, Subramanian, "Restart Perturbations for Lazy, Reversible Markov Chains: Trichotomy and Pre-cutoff Equivalence", arXiv 1907.02926
- Vial, Subramanian, "A Structural Result for Personalized PageRank and its Algorithmic Consequences", POMACS, June 2019
- Levin, Peres, Wilmer, "Markov Chains and Mixing Times"
- Diaconis, "The Cutoff Phenomenon in Finite Markov Chains", PNAS February 1996
- Basu, Hermon, Peres, "Characterization of cutoff for reversible Markov chains", SODA 2015

# References I


Aldous, David (1983). "Random walks on finite groups and rapidly mixing Markov chains". In: *Séminaire de Probabilités XVII 1981/82*. Springer, pp. 243–297.

Andersen, Reid, Fan Chung, Kevin Lang (2006). "Local graph partitioning using PageRank vectors". In: *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*. IEEE, pp. 475–486.

Andersen, Reid et al. (2008). "Local computation of PageRank contributions". In: *Internet Mathematics* 5.1-2, pp. 23–45.

Avena, Luca et al. (2018). "Random walks on dynamic configuration models: a trichotomy". In: *Stochastic Processes and their Applications*.

Avrachenkov, Konstantin et al. (2007). "Monte Carlo methods in PageRank computation: When one iteration is sufficient". In: *SIAM Journal on Numerical Analysis* 45.2, pp. 890–904.

Avrachenkov, Konstantin et al. (2015). "PageRank in undirected random graphs". In: *International Workshop on Algorithms and Models for the Web-Graph*. Springer, pp. 151–163.

Basu, Riddhipratim, Jonathan Hermon, Yuval Peres (2015). "Characterization of cutoff for reversible Markov chains". In: *Proceedings of the twenty-sixth annual ACM-SIAM symposium on Discrete algorithms*. Society for Industrial and Applied Mathematics, pp. 1774–1791.

Bayer, Dave, Persi Diaconis, et al. (1992). "Trailing the dovetail shuffle to its lair". In: *The Annals of Applied Probability* 2.2, pp. 294–313.

Ben-Hamou, Anna, Justin Salez, et al. (2017). "Cutoff for nonbacktracking random walks on sparse random graphs". In: *The Annals of Probability* 45.3, pp. 1752–1770.

# References II


Bordenave, Charles, Pietro Caputo, Justin Salez (2019). "Cutoff at the "entropic time" for sparse Markov chains". In: *Probability Theory and Related Fields* 173.1-2, pp. 261–292.

Caputo, Pietro, Matteo Quattropani (2019). "Mixing time of PageRank surfers on sparse random digraphs". In: *arXiv preprint arXiv:1905.04993.*

Chen, Ningyuan, Nelly Litvak, Mariana Olvera-Cravioto (2017). "Generalized PageRank on directed configuration networks". In: *Random Structures & Algorithms* 51.2, pp. 237–274.

Diaconis, Persi (1996). "The cutoff phenomenon in finite Markov chains". In: *Proceedings of the National Academy of Sciences* 93.4, pp. 1659–1664.

Diaconis, Persi, Mehrdad Shahshahani (1981). "Generating a random permutation with random transpositions". In: *Probability Theory and Related Fields* 57.2, pp. 159–179.

– (1987). "Time to reach stationarity in the Bernoulli–Laplace diffusion model". In: *SIAM Journal on Mathematical Analysis* 18.1, pp. 208–218.

Freschi, Valerio (2007). "Protein function prediction from interaction networks using a random walk ranking algorithm". In: *Bioinformatics and Bioengineering, 2007. BIBE 2007. Proceedings of the 7th IEEE International Conference on*. IEEE, pp. 42–48.

Gupta, Pankaj et al. (2013). "WTF: The who to follow service at twitter". In: *Proceedings of the 22nd international conference on World Wide Web*. ACM, pp. 505–514.

Jeh, Glen, Jennifer Widom (2003). "Scaling personalized web search". In: *Proceedings of the 12th international conference on World Wide Web*. ACM, pp. 271–279.

Kang, U, Charalampos E Tsourakakis, Christos Faloutsos (2009). "Pegasus: A peta-scale graph mining system implementation and observations". In: *Proceedings of the 2009 Ninth IEEE International Conference on Data Mining*. Washington, DC, USA, pp. 229–238.

Kloumann, Isabel M, Johan Ugander, Jon Kleinberg (2017). "Block models and personalized PageRank". In: *Proceedings of the National Academy of Sciences* 114.1, pp. 33–38.

Koutra, Danai, Joshua T Vogelstein, Christos Faloutsos (2013). "Deltacon: A principled massive-graph similarity function". In: *Proceedings of the 2013 SIAM International Conference on Data Mining*. SIAM, pp. 162–170.

Levin, David A, Yuval Peres, Elizabeth L Wilmer (2009). *Markov chains and mixing times*. American Mathematical Society.

Lofgren, Peter, Siddhartha Banerjee, Ashish Goel (2016). "Personalized PageRank estimation and search: A bidirectional approach". In: *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. ACM, pp. 163–172.

Lubetzky, Eyal, Allan Sly, et al. (2010). "Cutoff phenomena for random walks on random regular graphs". In: *Duke Mathematical Journal* 153.3, pp. 475–510.

Morrison, Julie L et al. (2005). "GeneRank: Using search engine technology for the analysis of microarray experiments". In: *BMC bioinformatics* 6.1, p. 233.

Page, Lawrence et al. (1999). *The PageRank citation ranking: Bringing order to the web.* Tech. rep. Stanford InfoLab.

Peres, Yuval, Perla Sousi (2015). "Mixing times are hitting times of large sets". In: *Journal of Theoretical Probability* 28.2, pp. 488–519.

# References IV


Tong, Hanghang, Christos Faloutsos, Jia-Yu Pan (2006). "Fast random walk with restart and its applications". In: *Sixth International Conference on Data Mining (ICDM'06)*. IEEE, pp. 613–622.

Vial, Daniel, Vijay Subramanian. "Restart perturbations for lazy, reversible Markov chains: trichotomy and pre-cutoff equivalence". In: *(In Preparation)* ().

– (2017). "On the role of clustering in Personalized PageRank estimation". In: *arXiv preprint arXiv:1706.01091*.

– (2019). "A Structural Result for Personalized PageRank and its Algorithmic Consequences". In: *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 3.2, p. 25.

Whang, Joyce Jiyoung et al. (2015). "Scalable data-driven pagerank: Algorithms, system issues, and lessons learned". In: *European Conference on Parallel Processing*. Springer, pp. 438–450.