Conductance and canonical paths for directed non-lazy walks

Ravi Montenegro

University of Massachusetts at Lowell

Random Structures & Algorithms - June 2011

1. Mixing Time of Markov Chains
   - Finite Markov Chains
   - Mixing Times
   - Spectral Gap and Canonical Paths

University of Massachusetts UMASS Lowell

Outline

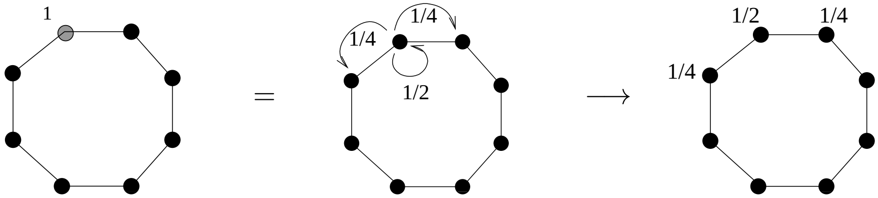
University of Massachusetts
UMASS Lowell

What is a Markov Chain?

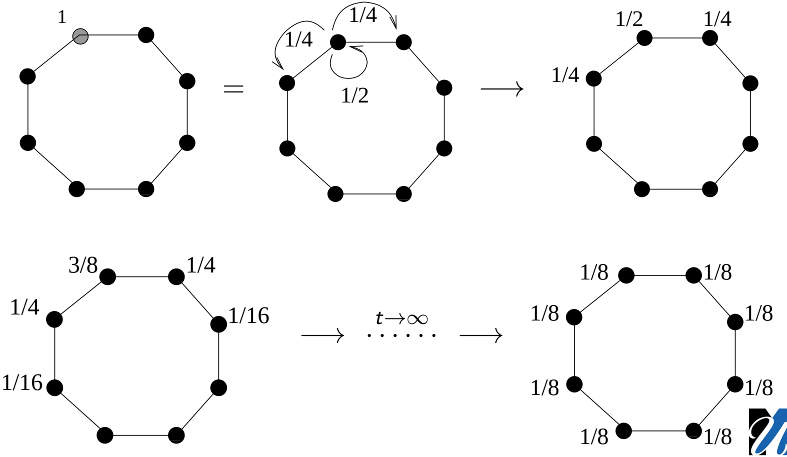
Finite Markov Chain

A random walk $X_0, X_1, X_2, \ldots$ on a graph which forgets its history.

$$P(X_{n+1} = y | X_n, X_{n-1}, \ldots, X_0) = P(X_{n+1} = y | X_n)$$



Lazy walk on a cycle



Markov chain terminology

- State space V.
- Transition matrix P.
- Distribution $\mathrm{p}^{(t)} = \mathrm{p}^{(0)}\mathrm{P}^t$
- Stationary distribution: $\pi\,\mathrm{P} = \pi$.

Theorem

Finite, irreducible, aperiodic $\Rightarrow \mathrm{p}^{(t)} \xrightarrow{t\to\infty} \pi$.

Markov chain terminology

- State space V.
- Transition matrix P.
- Distribution $\mathrm{p}^{(t)} = \mathrm{p}^{(0)}\mathrm{P}^t$
- Stationary distribution: $\pi\,\mathrm{P} = \pi$.

Theorem

Finite, irreducible, aperiodic $\Rightarrow \mathrm{p}^{(t)} \xrightarrow{t\to\infty} \pi$.

Applications of Markov chains

Method

Construct ergodic Markov chain which converges to distribution π.

Take some T-step distribution to be roughly stationary for the purposes of sampling from π.

- *Statistics:* statistical tests based on properties of random elements.
- *Statistical physics:* calculating expectations of random variables.
- *Computer science / combinatorial enumeration:* birthday attacks, approximate counting, volume estimation, integration, calculating permanent, etc..

Applications of Markov chains

Method

Construct ergodic Markov chain which converges to distribution π.

Take some T-step distribution to be roughly stationary for the purposes of sampling from π.

- *Statistics:* statistical tests based on properties of random elements.
- *Statistical physics:* calculating expectations of random variables.
- *Computer science / combinatorial enumeration:* birthday attacks, approximate counting, volume estimation, integration, calculating permanent, etc..

Example

Pollard Rho

Walk on $Z_n = \mathbb{Z}/n\mathbb{Z}$ with

$$\begin{aligned} i &\to i+1 \\ i &\to i+k \quad \text{(fixed constant } k\text{)} \\ i &\to 2i \end{aligned}$$

- Used for breaking Discrete Logarithm based codes.
- $\pi = U = \textit{uniform}$ if n is odd
- but not if n and k even

Question

How long until random walk draws good samples?

Example

Pollard Rho

Walk on $Z_n = \mathbb{Z}/n\mathbb{Z}$ with

$$\begin{aligned} i &\to i+1 \\ i &\to i+k \quad (\text{fixed constant } k) \\ i &\to 2i \end{aligned}$$

- Used for breaking Discrete Logarithm based codes.
- $\pi = U = \textit{uniform}$ if n is odd
- but not if n and k even

Question

How long until random walk draws good samples?

Mixing Time

Mixing Time

The mixing time $\tau(\epsilon)$ is the number of steps T required so that distance is less than ϵ.

Mixing time depends on choice of distance metric.

Relative density $f_t(v) = \frac{\mathrm{p}^{(t)}(v)}{\pi(v)}$

- *Variation Distance*: $\|\mathrm{p}^{(t)} - \pi\|_{TV} = \frac{1}{2}\sum_{v \in V} |\mathrm{p}^{(t)}(v) - \pi(v)| = \frac{1}{2}\|f_t - 1\|_{1,\pi}$
- *Chi-square distance*: $\|f_t - 1\|_{2,\pi}^2 = \sum_{v \in V} \pi(v)\left(\frac{\mathrm{p}^{(t)}(v)}{\pi(v)} - 1\right)^2$
- *Relative Pointwise distance*: $\max_v |f_t(v) - 1| = \|f_t - 1\|_{\infty,\pi}$

Mixing Time

Mixing Time

The mixing time $\tau(\epsilon)$ is the number of steps T required so that distance is less than ϵ.

Mixing time depends on choice of distance metric.

Relative density $f_t(v) = \frac{\mathrm{p}^{(t)}(v)}{\pi(v)}$

- *Variation Distance*:
$$\|\mathrm{p}^{(t)} - \pi\|_{TV} = \frac{1}{2} \sum_{v \in V} |\mathrm{p}^{(t)}(v) - \pi(v)| = \frac{1}{2} \|f_t - 1\|_{1,\pi}$$
- *Chi-square distance*: $\|f_t - 1\|_{2,\pi}^2 = \sum_{v \in V} \pi(v) \left(\frac{\mathrm{p}^{(t)}(v)}{\pi(v)} - 1 \right)^2$
- *Relative Pointwise distance*: $\max_v |f_t(v) - 1| = \|f_t - 1\|_{\infty,\pi}$

Methods for Bounding Mixing Times

Total Variation

Direct Computation, Coupling, Strong Stationary Time.

Chi-Square

Direct Computation, Spectral Gap, Canonical Paths, Comparison, Fourier Analysis.

Methods for Bounding Mixing Times

Total Variation

Direct Computation, Coupling, Strong Stationary Time.

Chi-Square

Direct Computation, Spectral Gap, Canonical Paths, Comparison, Fourier Analysis.

Cayley Graph

Finite group G.

Generating set S.

$\forall g \in G \ \exists s_1, s_2, \dots : s_1 s_2 \cdots s_n = g$ (inverses not in S)

Random Walk: Given prob distn $p : S \to [0,1]$ take

$$P(g,h) = p(g^{-1}h)$$

ie. choose random generator $s \in S$ and transition

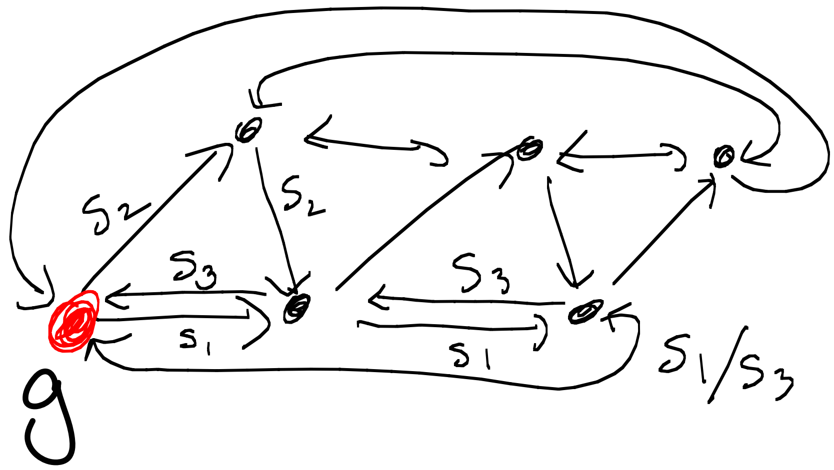
$$g \to gs$$

Example

$s_2^2 = s_1$

$s_1^3 = id$

$s_3 = s_1^{-1}$



Question: How long does this take to mix?

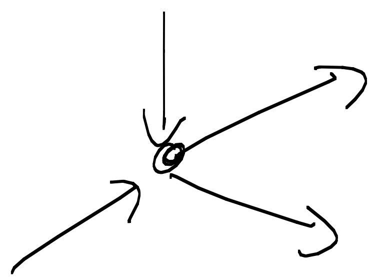
i.e. generate near-uniform group element.

If Δ = diameter then

Babai: $O\left(\frac{\Delta^2}{\min p(s)} \log \frac{|G|}{\varepsilon}\right)$ if $\underbrace{S = S^{-1}}_{\text{reversible}}$ & $\underbrace{p(id) \geq \frac{1}{2}}_{\text{lazy}}$

Diaconis & Saloff-Coste: Ditto if merely $p(id) \geq \frac{1}{2}$, or $S = S^{-1}$ and $id \in S$.

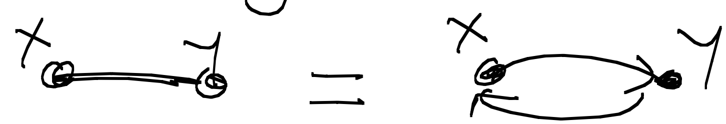
Eulerian Graphs



$\deg_{in}(x) = \deg_{out}(x)$ everywhere

$\Rightarrow \pi(x) \propto \deg(x)$

Ex: Undirected graph



Fact: Max-deg walk mixes in $O\left(n^2 d \log \frac{1}{\varepsilon}\right)$ @ $n = |V|$
if lazy $\left(P(x,x) \geq \frac{1}{2}\right)$ $d = \max \deg(x)$

Question: What if non-lazy?

Non-lazy and non-reversible

Cayley: $id \in S$ (not lazy!): $O\left(\frac{\Delta^2}{\min p(s)} \log \frac{|G|}{\Delta \varepsilon}\right)$

$id \notin S$: Take Δ along odd length path of $p \to p^* \to p \to p^* \to \cdots \to p$

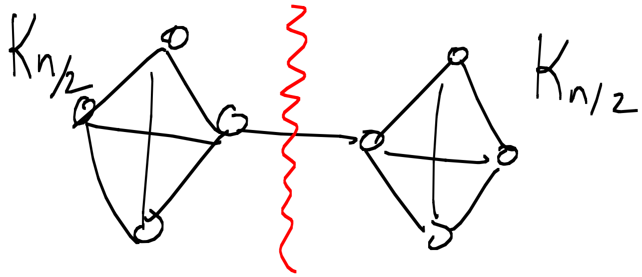
p^* = reversal = walk backwards on edges.

Eulerian: Self loop @ each vertex: $O(n^2 d \log \frac{1}{\varepsilon})$

Non looping: Same if connected by odd length paths of $p \to p^* \to p \to p^* \to \cdots \to p$

Method: Extend canonical paths to non-reversible non-lazy case.

Conductance

A  A^c

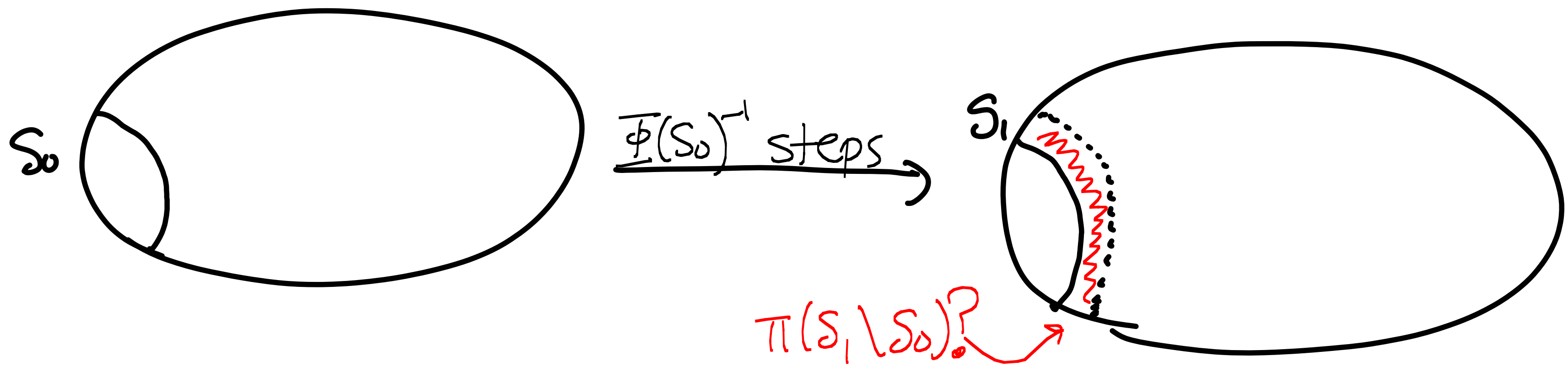
Bottleneck @ A, i.e. hard for walk to go from A to A^c.

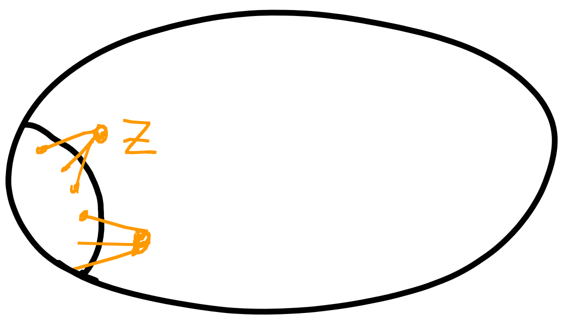
Definition: If $Q(A,B) := \sum_{\substack{x\in A\\ y\in B}} \pi(x)P(x,y)$ then

$$\Phi(A) = \frac{Q(A,A^c)}{\pi(A)} \left(= \Pr[X_1 \in A^c \mid X_0 \in_\pi A] \right)$$

Theorem [JS, LS]: Mixing time $O\left(\frac{1}{\Phi^2}\log\frac{1}{\pi_0 \varepsilon}\right)$ if $\Phi = \min_{\pi(A)\le\frac{1}{2}} \Phi(A)$

Intuition for Conductance





$$\sum_{x\in A} \pi(x)P(x,z) \le \frac{1}{2}\pi(z)$$

$$\Rightarrow \frac{1}{2}\pi(S_1 \setminus S_0) \ge Q(S_0, S_0^c)$$

$$\boxed{\begin{aligned} \Rightarrow \pi(S_1) &\ge \pi(S_0) + 2Q(S_0, S_0^c) \\ &= \pi(S_0)\left(1 + 2\Phi(S_0)\right) \end{aligned}}$$

$$\pi(S_0) \longrightarrow \pi(S_0)(1+\Phi) \longrightarrow \overset{t \text{ rounds}}{\cdots} \longrightarrow \pi(S_0)(1+\Phi)^t$$

Reach

by

$$T = \frac{\log \frac{1/2}{\pi(S_0)}}{\log(1+\Phi)} \approx \frac{1}{\Phi} \log \frac{1}{2\pi(S_0)}$$

$\vdots$

T more rounds

$\longrightarrow$ Mixing time $2T\Phi^{-1} \sim \frac{2}{\Phi^2} \log \frac{1}{2\pi(x)}$

Conductance Profile

$$\frac{1}{\Phi(S_0)} + \frac{1}{\Phi(S_1)} + \cdots + \frac{1}{\Phi(S_t)} \le \frac{1}{\Phi(S_0)} \cdot \frac{\pi(S_1)-\pi(S_0)}{Q(S_0, S_0^c)} + \cdots$$

$$= \frac{1}{\pi(S_0)\,\Phi(S_0)^2}\left(\pi(S_1)-\pi(S_0)\right) + \cdots$$

Conductance Profile: $\Phi(r) = \min_{\pi(A) \le r} \Phi(A)$

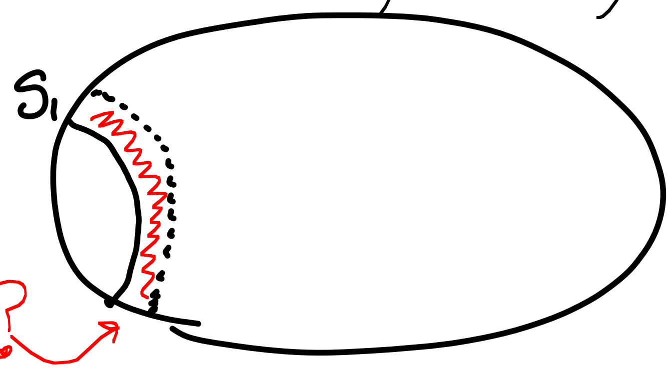
Let $x_i = \pi(S_i)$:

$$\int_{x_0}^{x_1} \frac{1}{x\,\Phi(x)^2}\,dx + \int_{x_1}^{x_2} \frac{1}{x\,\Phi(x)^2}\,dx + \cdots$$

[LK/MP/GMT]: Mixing in $\displaystyle\int_{\pi_0}^{1/2} \frac{dx}{x\,\Phi(x)^2}$

Vertex Expansion

Room to improve: Directly study $\pi(S_1 \setminus S_0)$



$\pi(S_1 \setminus S_0)$?

KLM: Blocking Conductance ← block a bit, see what gets by

MP: Evolving Sets

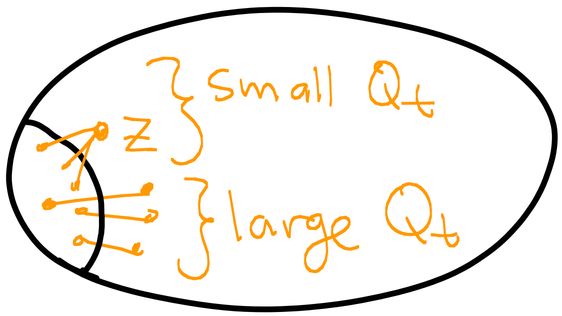
Recall: Lazy $\Rightarrow$ $\frac{1}{2}\pi(S_1 \setminus S_0) \geq Q(S_0, S_0^c)$

New: Threshold Conductance

$$Q_t(A, B) = \sum_{y \in B} \min\{Q(A, y), t\pi(y)\}$$

$\rightarrow t\pi(S_1 \setminus S_0) \geq Q_t(A, A^c) \rightarrow \pi(S_1 \setminus S_0) \geq Q_t(A, A^c)/t$

$$\pi(S_1 \setminus S_0) \ge Q_t(A, A^c)/t$$



$$\frac{1}{\Phi_t(S_0)} + \frac{1}{\Phi_t(S_1)} + \dots + \frac{1}{\Phi_t(S_t)} \le \frac{1}{\Phi_t(S_0)} \cdot \frac{\pi(S_1)-\pi(S_0)}{Q_t(S_0, S_0^c)/t} + \dots$$

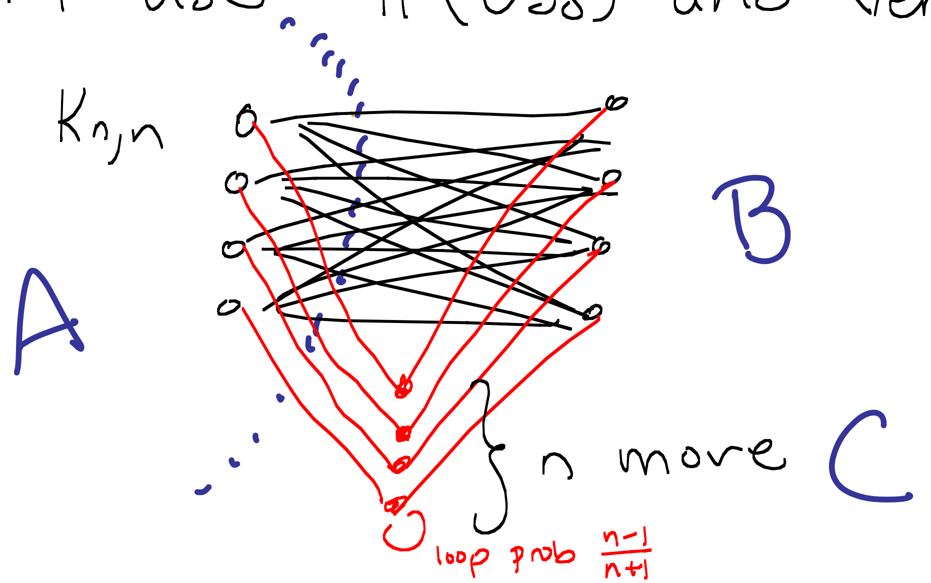
$$= \frac{t}{\pi(S_0)\,\Phi_t(S_0)^2}\left(\pi(S_1)-\pi(S_0)\right) + \dots$$

Threshold Conductance: $\Phi_t(r) = \min\limits_{\pi(A)\le r} \Phi_t(A) = \min \frac{Q_t}{\pi}$

et $x_i = \pi(S_i)$:

$$\int_{x_0}^{x_1} \frac{t}{x\,\Phi_t(x)^2}\,dx + \int_{x_1}^{x_2} \frac{t}{x\,\Phi_t(x)^2}\,dx + \dots \sim \boxed{\int_{\pi_0}^{1/2} \frac{t}{x\,\Phi_t(x)^2}\,dx}$$

Why not use $\pi(\partial S_0)$ and vertex expansion?



Expected growth $|A|=n \longrightarrow \cancel{n+n}$?

$|A|=n \longrightarrow n+1$

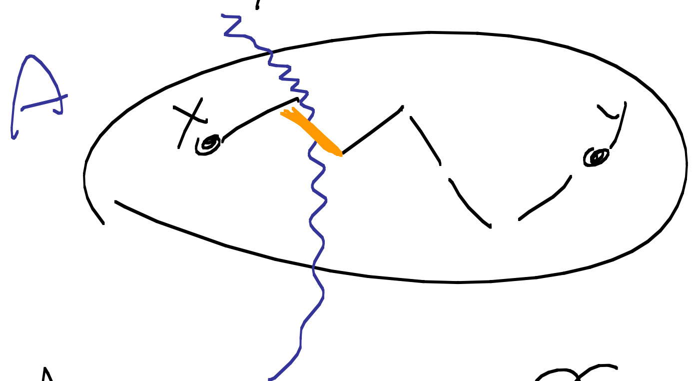
<u>Mixing Time</u>

$$Q_{\frac{1}{n+1}}(C,C^c) = \pi(C)\frac{2}{n+1} \longrightarrow \Phi_{1/n+1}(C) = \frac{2}{n+1}$$

$$\Rightarrow \text{mixing } \int_{1/3n}^{1/2} \frac{1/(n+1)}{x/(n+1)^2}\,dx = (n+1)\log 3n$$

Canonical Paths

For each $x,y \in V$ construct path γ_{xy} from x to y



If $x \in A$, $y \in A^c$ $\longrightarrow$ γ_{xy} includes edge from $A \to A^c$

Let

$$\rho_e = \max_{e=(a,b)\in E} \frac{1}{\pi(a)P(a,b)} \sum_{\gamma_{xy} \ni (a,b)} \pi(x)\pi(y)$$

Then

$$\sum_{\substack{x\in A \\ y \in A^c}} \pi(x)\pi(y) = \pi(A)\pi(A^c) \quad \& \quad \leq \rho_e \cdot \pi(a)P(a,b) \text{ per edge}$$

$$\Rightarrow Q(A,A^c) \geq \sum \pi(a)P(a,b) \geq \frac{\pi(A)\pi(A^c)}{\rho_e}$$

$Q(A,A^c) \geq \sum \pi(a)P(a,b) \geq \frac{\pi(A)\pi(A^c)}{\rho_e} \Rightarrow \Phi \geq 1/\rho_e$

$\boxed{\text{mix in } \rho_e^2 \log \frac{1}{\pi_0}}$

[DS, Sin] Mix in $\boxed{\rho_e \, \ell \log \frac{1}{\pi_0}}$

ℓ ← max path length

New: Vertex congestion

$$\rho_v = \max_{v \in V} \frac{1}{\pi(v)} \sum_{\gamma_{xy} \ni v} \pi(x)\pi(y) \quad (\leq \rho_e)$$

Then

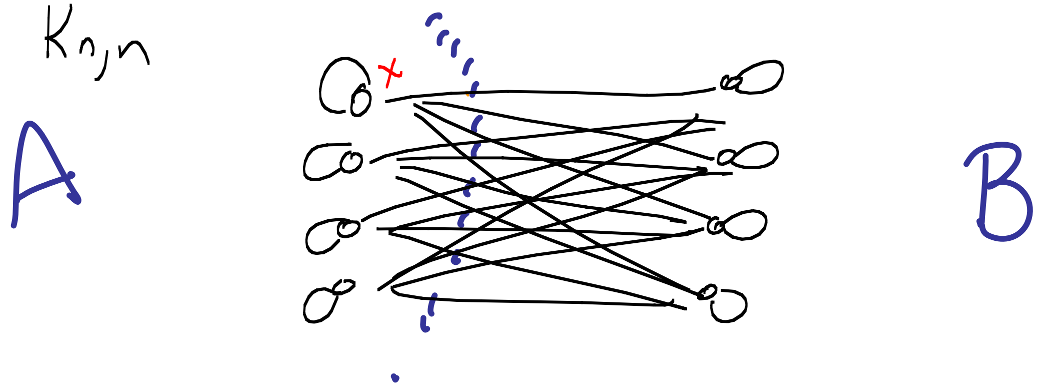
$$\Phi_{\rho_v/\rho_e} \geq 1/\rho_e \quad (\text{versus } \Phi \geq 1/\rho_e)$$

$$\Rightarrow \text{mix in } \int \frac{\rho_v/\rho_e}{x\,(1/\rho_e)^2}\,dx = \boxed{\rho_e \rho_v \log \frac{1}{\pi_0}}$$

Non-lazy

Holding probability: $\alpha = \min\limits_{v \in V} P(v,v) < \frac{1}{2}$

Example: $K_{n,n}$



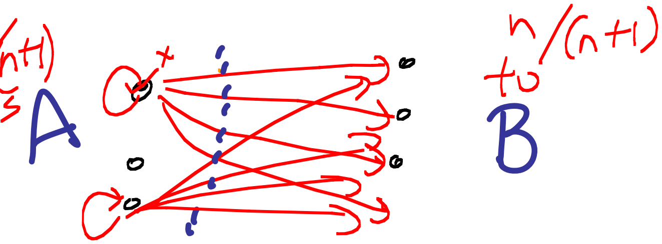
w/o loops: $x \longrightarrow B \longrightarrow A \longrightarrow B \longrightarrow \cdots$ Never mixes!

Problem: Move from B to $\pi(A)=\pi(B)=\frac{1}{2}$. Stuck @ $\frac{1}{2}$ space.

with loops: $S_0 \longrightarrow B+\frac{1}{n}S_0 \longrightarrow (1-\frac{1}{n})A+\frac{1}{n}B \longrightarrow (1-\frac{1}{n})^2 B + \frac{2}{n}A \longrightarrow \cdots$ n rounds

New: Move from B to $\pi(A)=\pi(B)=\frac{1}{2}$ but a little escapes.

How much escapes $\frac{1}{2}$ space?



Modified flow: $\Psi(A) = \min_{\pi(B)=\pi(A)} Q(A, B^c)$

Modified conductance: $\phi(A) = \dfrac{\Psi(A)}{\pi(A)}$

Same intuition when non-lazy

$\Rightarrow$ mix in $\displaystyle\int \frac{1}{x\,\phi(x)^2}\,dx$

Threshold: $\Psi_t(A) = \min_{\pi(B)=\pi(A^c)} Q_t(A, B^c)$ & ϕ_t

$\Rightarrow$ mix in $\displaystyle\int \frac{t}{x\,\phi_t(x)^2}\,dx$

$\rightarrow$ If $t \le \alpha$ then $\Psi_t(A) = Q_t(A, A^c)$ & get $\displaystyle\int \frac{t}{x\,\Phi_t(x)^2}\,dx$

Canonical Paths: Non-Lazy

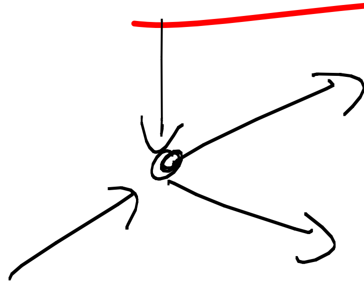
Recall $\Phi_{\rho_v/\rho_e} \geq 1/\rho_e$.

If $\frac{\rho_v}{\rho_e} \leq \alpha$ then still mix in $\rho_e \rho_v \log \frac{1}{\pi_0}$

If $\frac{\rho_v}{\rho_e} > \alpha$ then $\Phi_\alpha \geq \frac{\alpha}{\rho_v/\rho_e} \Phi_{\rho_v/\rho_e} \geq \frac{\alpha}{\rho_v} \Rightarrow$ mix $\frac{\rho_v^2}{\alpha} \log \frac{1}{\pi_0}$

Mix in $\rho_v \max\left\{\frac{\rho_v}{\alpha}, \rho_e\right\} \log \frac{1}{\pi_0}$

Eulerian max-degree (self-loops)



$\deg_{in}(x) = \deg_{out}(x)$ everywhere

$\Rightarrow \pi(x) \propto \deg(x)$

n vertices, max degree d.

- $\alpha \geq \frac{1}{d}$
- At worst every path goes through every vertex:
  - $$\rho_e = \max_{e=(a,b)\in E} \frac{1}{\pi(a)P(a,b)} \sum_{\gamma_{xy} \ni (a,b)} \pi(x)\pi(y)$$
    $$\leq \frac{1}{\frac{1}{n}\frac{1}{d}} \cdot 1 = nd$$
  - $$\rho_v = \max_{v \in V} \frac{1}{\pi(v)} \sum_{\gamma_{xy} \ni v} \pi(x)\pi(y) \leq \frac{1}{\frac{1}{n}} \cdot 1 = n$$

$$\Rightarrow \text{mix in } \rho_v \max\left\{\frac{\rho_v}{\alpha}, \rho_e\right\} \log \frac{1}{\pi_0 \rho_v} \leq \underline{\underline{n^2 d}}$$

Cayley Graph (ideS)

$G = \langle S \rangle$, inverses not in S

$$P(g,h) = p(g^{-1}h)$$

If $x,y \in G$ then $\gamma_{id,x^{-1}y}$ induces path $x \to y$

$\Rightarrow \rho_v$ same for every vertex

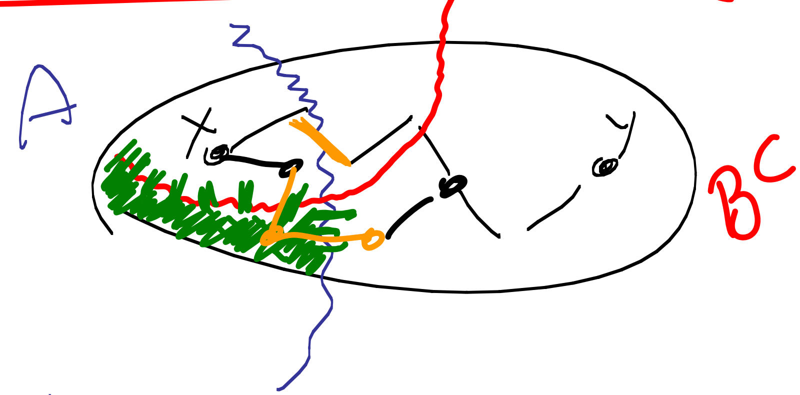
$$\rho_v = \max_{v \in V} \frac{1}{\pi(v)} \sum_{\gamma_{xy} \ni v} \pi(x)\pi(y) \le \frac{1}{1/|G|} \cdot \frac{1}{|G|} \Big(\underbrace{|G|^2 \Delta}_{\text{all edges}} \cdot \frac{1}{|G|} \frac{1}{|G|} \Big) = \Delta$$

$$\rho_e = \max_{e=(a,b) \in E} \frac{1}{\pi(a)P(a,b)} \sum_{\gamma_{xy} \ni (a,b)} \pi(x)\pi(y) \le \frac{\rho_v}{\min p(s)}$$

$$\alpha \ge \min p(s)$$

$\Rightarrow$ mix in $\rho_v \max\{\frac{\rho_v}{\alpha}, \rho_e\} \log \frac{1}{\pi_0 \rho_v} \le \frac{\Delta^2}{\min p(s)} \log \frac{|G|}{\Delta}$

Canonical Paths ($\alpha \sim 0$)

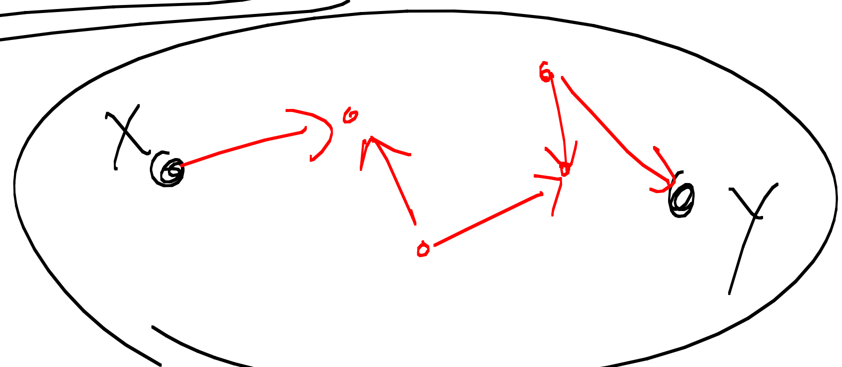


Threshold: $\psi_t(A) = \min\limits_{\pi(B)=\pi(A^c)} Q_t(A, B^c) \geq Q_t(A, V) - t\pi(A)$

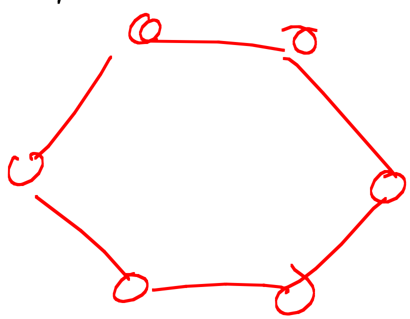
Lemma: $\phi^{P_0^*}(A) \geq \dfrac{P_0^*}{2P_V}$ where $P_0^* = \min P^*(x,y)$
$= \min \pi(y)P(y,x)/\pi(x)$

and paths odd length alternating between edges of P and P^*.

Paths of interest

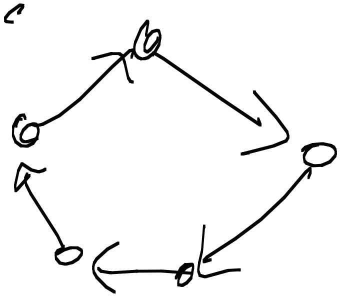


Why parity?



Doesn't mix

Why P & P*?



Doesn't mix

Cayley Graph ($\alpha \sim 0$)

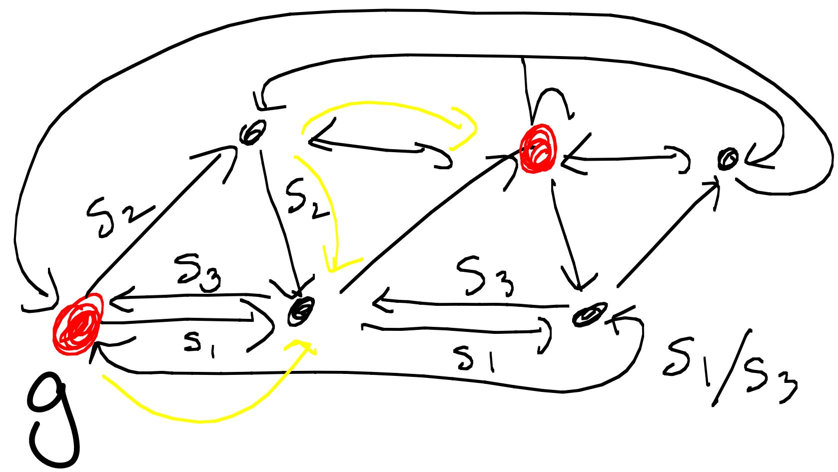
$G = \langle S \rangle$

Example

$s_2^2 = s_1$

$s_1^3 = id$

$s_3 = s_1^{-1}$



If $id \notin S$ or $p(id)$ small let Δ^* be diameter among odd length paths alternating between edges of $P \to P^* \to \cdots \to P$.

Then mixes in $O\left(\frac{\Delta^{*2}}{\min p(s)} \log \frac{|G|}{\Delta^*}\right)$