Mixing times and Coarse Ricci curvature on the permutation group

Batı Şengül

with Nathanaël Berestycki

University of Cambridge/University of Bath

June 2014

$\mathcal{S}_n$ is the set of permutations of $\{1, \ldots, n\}$.

$\mathcal{S}_n$ is the set of permutations of $\{1, \dots, n\}$.
$X = (X_t : t \geq 0)$ continuous time random walk: $X_t = \gamma_1 \circ \cdots \circ \gamma_{N_t}$ with $\gamma_i \overset{\text{i.i.d.}}{\sim} \gamma$

- Transposition: $\gamma = (i, j)$ with $i \neq j$ both uniform

$\mathcal{S}_n$ is the set of permutations of $\{1, \ldots, n\}$.
$X = (X_t : t \geq 0)$ continuous time random walk: $X_t = \gamma_1 \circ \cdots \circ \gamma_{N_t}$ with $\gamma_i \overset{\text{i.i.d.}}{\sim} \gamma$

- Transposition: $\gamma = (i, j)$ with $i \neq j$ both uniform
- k-cycles: $\gamma = (x_1, \ldots, x_k)$ with $x_i \neq x_j$ for all $i \neq j$ chosen uniformly

$\mathcal{S}_n$ is the set of permutations of $\{1, \ldots, n\}$.
$X = (X_t : t \geq 0)$ continuous time random walk: $X_t = \gamma_1 \circ \cdots \circ \gamma_{N_t}$ with $\gamma_i \overset{\text{i.i.d.}}{\sim} \gamma$

- Transposition: $\gamma = (i, j)$ with $i \neq j$ both uniform
- k-cycles: $\gamma = (x_1, \ldots, x_k)$ with $x_i \neq x_j$ for all $i \neq j$ chosen uniformly

X has invariant distribution μ (when k is even μ is uniform on $\mathcal{S}_n$)

$$d_{TV}(t) = \|X_t - \mu\|_{TV} = \sum_{\sigma \in \mathcal{S}_n} |\mathbb{P}(X_t = \sigma) - \mu(\sigma)|$$
$$= \inf_{X_t' \sim X_t, Y' \sim \mu} \mathbb{P}(X_t' \neq Y')$$

$$d_{TV}(t) = \|X_t - \mu\|_{TV} = \sum_{\sigma \in \mathcal{S}_n} |\mathbb{P}(X_t = \sigma) - \mu(\sigma)|$$
$$= \inf_{X'_t \sim X_t, Y' \sim \mu} \mathbb{P}(X'_t \neq Y')$$

Theorem (Diaconis & Shahshahani (1981))

Let $t_{\text{mix}} = (1/2)n \log n$. Then each $\epsilon > 0$ we have that for the **transposition random walk** *$(k = 2)$:*

$$\lim_{n\to\infty} d_{TV}((1-\epsilon)\, t_{\text{mix}}) = 1 \qquad \lim_{n\to\infty} d_{TV}((1+\epsilon)\, t_{\text{mix}}) = 0.$$

$$d_{TV}(t) = \|X_t - \mu\|_{TV} = \sum_{\sigma \in \mathcal{S}_n} |\mathbb{P}(X_t = \sigma) - \mu(\sigma)|$$
$$= \inf_{X_t' \sim X_t, Y' \sim \mu} \mathbb{P}(X_t' \neq Y')$$

Theorem (Diaconis & Shahshahani (1981))


Let $t_{\text{mix}} = (1/2) n \log n$. *Then each* $\epsilon > 0$ *we have that for the* **transposition random walk** *(*$k = 2$*):*

$$\lim_{n\to\infty} d_{TV}((1-\epsilon)\, t_{\text{mix}}) = 1 \qquad \lim_{n\to\infty} d_{TV}((1+\epsilon)\, t_{\text{mix}}) = 0.$$

The above is referred to as the **cut-off phenomenon** and t_{mix} is called the **mixing time**.

Theorem (Berestycki & Ş. (2014))

Let $t_{\text{mix}} = (1/k)n \log n$ *and suppose that* $k = k(n) = o(n)$. *Then each* $\epsilon > 0$ *we have that*

$$\lim_{n\to\infty} d_{TV}((1-\epsilon)\,t_{\text{mix}}) = 1 \qquad \lim_{n\to\infty} d_{TV}((1+\epsilon)\,t_{\text{mix}}) = 0.$$

Theorem (Lulov (1996))

Cutoff phenomenon + $t_{\text{mix}} = (1/k) n \log n$ *for* $k \geq n/2$.

Theorem (Lulov (1996))

Cutoff phenomenon $+\ t_{\text{mix}} = (1/k)n \log n$ *for* $k \geq n/2$.

Theorem (Roichman(1996))

$t_{\text{mix}} \leq (C/k)n \log n$ *when* $k = o(n)$.

Theorem (Lulov (1996))

Cutoff phenomenon $+ \ t_{\text{mix}} = (1/k)n\log n$ *for* $k \geq n/2$.

Theorem (Roichman(1996))

$t_{\text{mix}} \leq (C/k)n\log n$ *when* $k = o(n)$.

Theorem (Roussel (2000))

Cutoff phenomenon $+ \ t_{\text{mix}} = (1/k)n\log n$ *for* $k \leq 6$.

Theorem (Lulov (1996))

Cutoff phenomenon + $t_{\text{mix}} = (1/k)n \log n$ *for* $k \geq n/2$.

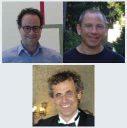
Theorem (Roichman(1996))

$t_{\text{mix}} \leq (C/k)n \log n$ *when* $k = o(n)$.

Theorem (Roussel (2000))

Cutoff phenomenon + $t_{\text{mix}} = (1/k)n \log n$ *for* $k \leq 6$.

Theorem (Berestycki, Schramm, Zeitouni (2011))



Cutoff phenomenon + $t_{\text{mix}} = (1/k)n \log n$ *for* k **fixed**

L^1 transportation distance:

$$W_1(\nu, \pi) = \inf_{X \sim \nu, Y \sim \pi} \mathbb{E}[d(X, Y)]$$

L^1 transportation distance:

$$W_1(\nu, \pi) = \inf_{X \sim \nu, Y \sim \pi} \mathbb{E}[d(X, Y)]$$

Definition (Ollivier (2009))

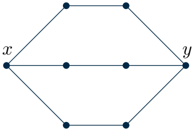
Let (E, d) be a graph and $\{m_x\}_{x \in E}$ a collection of probability measures on E. For $x, y \in E$ with $x \neq y$ define

$$\kappa(x, y) = 1 - \frac{W_1(m_x, m_y)}{d(x, y)}.$$

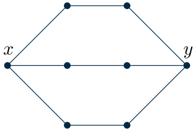
Define the **coarse Ricci curvature** *as* $\kappa = \inf_{x \neq y} \kappa(x, y)$.

$\kappa(x, y) > 0$ if and only if $W_1(m_x, m_y) < d(x, y)$:

$\kappa(x, y) > 0$ if and only if $W_1(m_x, m_y) < d(x, y)$:

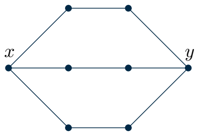


$\kappa(x, y) > 0$ if and only if $W_1(m_x, m_y) < d(x, y)$:

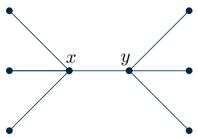


$\kappa(x, y) < 0$ if and only if $W_1(m_x, m_y) > d(x, y)$:

$\kappa(x, y) > 0$ if and only if $W_1(m_x, m_y) < d(x, y)$:



$\kappa(x, y) < 0$ if and only if $W_1(m_x, m_y) > d(x, y)$:



Now $(E, d) = (\mathcal{S}_n, d)$ where $d(\sigma, \sigma')$ minimum number of transpositions $\tau_1, \ldots, \tau_n$ so that $\sigma = \sigma' \circ \tau_1 \circ \cdots \circ \tau_n$.

Now $(E, d) = (\mathcal{S}_n, d)$ where $d(\sigma, \sigma')$ minimum number of transpositions $\tau_1, \ldots, \tau_n$ so that $\sigma = \sigma' \circ \tau_1 \circ \cdots \circ \tau_n$. For $c > 0$ we take $m_x = X^x_{cn/2}$ **transposition random walk** started at x.

Now $(E, d) = (\mathcal{S}_n, d)$ where $d(\sigma, \sigma')$ minimum number of transpositions $\tau_1, \ldots, \tau_n$ so that $\sigma = \sigma' \circ \tau_1 \circ \cdots \circ \tau_n$. For $c > 0$ we take $m_x = X^x_{cn/2}$ **transposition random walk** started at x.

Theorem (Berestycki & Ş. (2014))

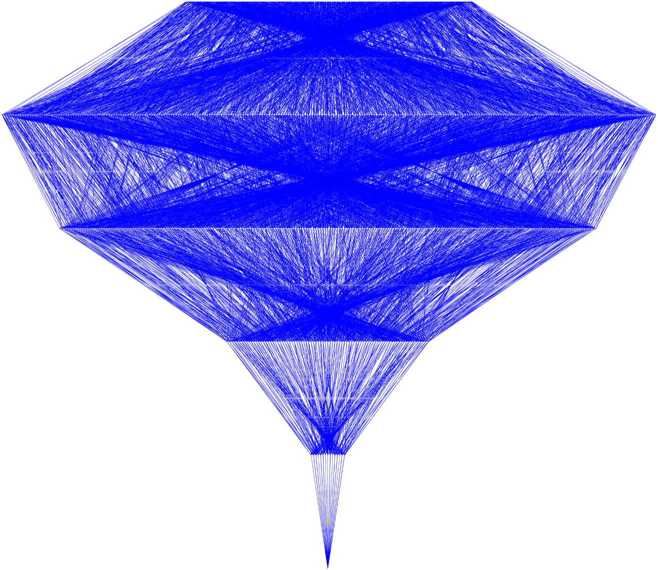
For $c > 0$ let κ_c denote the coarse Ricci curvature. Then for $c \leq 1$

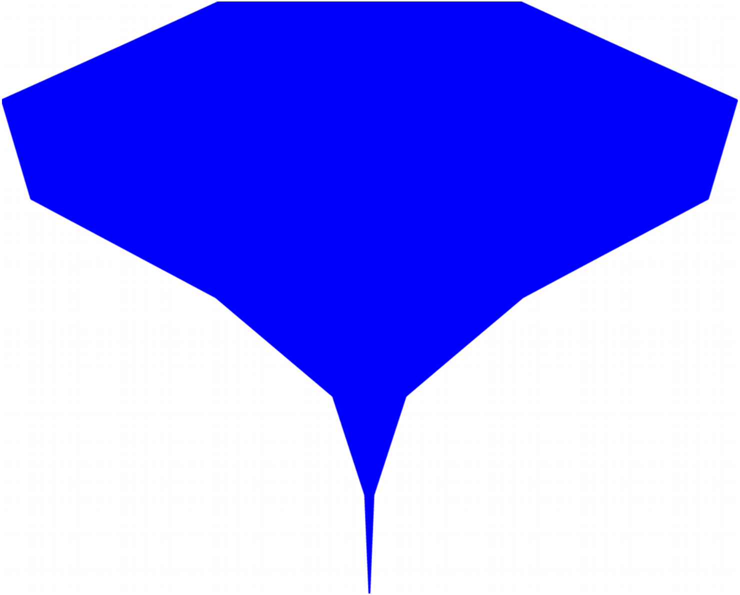
$$\lim_{n\to\infty} \kappa_c = 0$$

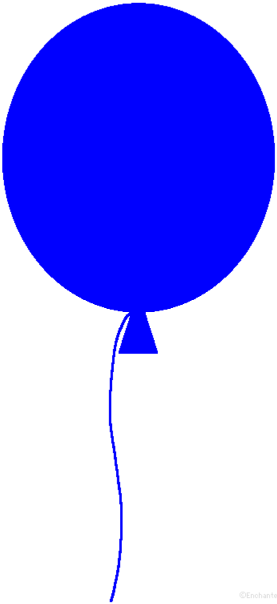
and for $c > 1$

$$\theta(c)^4 \leq \liminf_{n\to\infty} \kappa_c \leq \limsup_{n\to\infty} \kappa_c \leq \theta(c)^2$$

where $\theta(c) \in (0,1)$ is the solution to $\theta(c) = 1 - e^{-c\theta(c)}$.







Proof of Mixing

$$||\nu - \pi||_{TV} = \inf_{X \sim \nu, Y \sim \pi} \mathbb{E}[\mathbf{1}_{\{X \neq Y\}}] \leq \inf_{X \sim \nu, Y \sim \pi} \mathbb{E}[d(X, Y)] = W_1(\nu, \pi)$$

Proof of Mixing

$$\|\nu - \pi\|_{TV} = \inf_{X\sim\nu, Y\sim\pi} \mathbb{E}[\mathbf{1}_{\{X\neq Y\}}] \leq \inf_{X\sim\nu, Y\sim\pi} \mathbb{E}[d(X,Y)] = W_1(\nu,\pi)$$

Proposition (Bubley & Dyer (1997), Ollivier (2009))

For each $s \geq 0$,

$$\|m_x^{*s} - \mu\|_{TV} \leq diam(E)(1-\kappa)^s$$

Proof of Mixing

For $c > 1$ we have that

$$\lim_{n\to\infty} d_{TV}\left(\frac{c}{-\log(1-\theta(c)^4)}\frac{n}{2}\log n\right) = 0.$$

Proof of Mixing

For $c > 1$ we have that

$$\lim_{n\to\infty} d_{TV}\left(\frac{c}{-\log(1-\theta(c)^4)}\frac{n}{2}\log n\right) = 0.$$

Easy computation:

$$\lim_{c\uparrow\infty}\frac{c}{-\log(1-\theta(c)^4)} = 1$$

Proof of Mixing

For $c > 1$ we have that

$$\lim_{n\to\infty} d_{TV}\left(\frac{c}{-\log(1-\theta(c)^4)}\frac{n}{2}\log n\right) = 0.$$

Easy computation:

$$\lim_{c\uparrow\infty} \frac{c}{-\log(1-\theta(c)^4)} = 1$$

Hence for each $\epsilon > 0$,

$$\lim_{n\to\infty} d_{TV}((1+\epsilon)(1/2)n\log n) = 0$$

Let $t = cn/2$, then

$$\kappa_c(x, y) = 1 - \frac{\inf_{X'_t \sim X^x_t, Y'_t \sim X^y_t} \mathbb{E}[d(X'_t, Y'_t)]}{d(x, y)}$$

Let $t = cn/2$, then

$$\kappa_c(\mathrm{id}, x) = 1 - \frac{\inf_{X'_t \sim X^{\mathrm{id}}_t, Y'_t \sim X^x_t} \mathbb{E}[d(X'_t, Y'_t)]}{d(id, x)}$$

Let $t = cn/2$, then

$$\kappa_c(\mathrm{id}, \tau) = 1 - \frac{\inf_{X'_t \sim X^{\mathrm{id}}_t, Y'_t \sim X^{\tau}_t} \mathbb{E}[d(X'_t, Y'_t)]}{1}$$

Let $t = cn/2$, then

$$\kappa_c(\text{id}, \tau) = 1 - \frac{\inf_{X'_t \sim X^{\text{id}}_t, Y'_t \sim X^{\tau}_t} \mathbb{E}[d(X'_t, Y'_t)]}{1}$$

Take τ to be a uniform transposition and take $\tau_1 = \tau$, coupling up to time $t_0 = t - \log n$

$$X^{\text{id}}_{t_0} = \tau_1 \circ \cdots \circ \tau_{N_{t_0}}$$
$$X^{\tau}_{t_0} = \tau_1 \circ \cdots \circ \tau_{N_{t_0}} \circ \tau_{N_{t_0}+1}.$$

Let $t = cn/2$, then

$$\kappa_c(\mathrm{id}, \tau) = 1 - \frac{\inf_{X'_t \sim X^{\mathrm{id}}_t, Y'_t \sim X^{\tau}_t} \mathbb{E}[d(X'_t, Y'_t)]}{1}$$

Take τ to be a uniform transposition and take $\tau_1 = \tau$, coupling up to time $t_0 = t - \log n$

$$\begin{aligned} X^{\mathrm{id}}_{t_0} &= \tau_1 \circ \cdots \circ \tau_{N_{t_0}} \\ X^{\tau}_{t_0} &= \tau_1 \circ \cdots \circ \tau_{N_{t_0}} \circ \tau_{N_{t_0}+1}. \end{aligned}$$

Theorem (Schramm (2005))

There exists a set $A(t_0) \subset \{1, \ldots, n\}$ of size $\approx \theta(c)n$ such that X_{t_0} restricted to $A(t_0)$ "looks like" the uniform permutation.

Two cases: $\tau_{N_{t_0}+1} = (i,j)$ with $i,j \in A(t_0)$. Then there exists a coupling during $[t_0, t]$ such that

$$X_t^{\mathrm{id}} = X_t^{\tau}.$$

Two cases: $\tau_{N_{t_0}+1} = (i, j)$ with $i, j \in A(t_0)$. Then there exists a coupling during $[t_0, t]$ such that

$$X_t^{\text{id}} = X_t^{\tau}.$$

$\tau_{N_{t_0}+1} = (i, j)$ with $i \notin A(t_0)$ or $j \notin A(t_0)$. Then take

$$\begin{aligned} X_t^{\text{id}} &= \tau_1 \circ \cdots \circ \tau_{N_t} \\ X_t^{\tau} &= \tau_1 \circ \cdots \circ \tau_{N_t} \circ \tau_{N_t+1}. \end{aligned}$$

Two cases: $\tau_{N_{t_0}+1} = (i,j)$ with $i,j \in A(t_0)$. Then there exists a coupling during $[t_0, t]$ such that

$$X_t^{\mathrm{id}} = X_t^{\tau}.$$

$\tau_{N_{t_0}+1} = (i,j)$ with $i \notin A(t_0)$ or $j \notin A(t_0)$. Then take

$$\begin{aligned} X_t^{\mathrm{id}} &= \tau_1 \circ \cdots \circ \tau_{N_t} \\ X_t^{\tau} &= \tau_1 \circ \cdots \circ \tau_{N_t} \circ \tau_{N_t+1}. \end{aligned}$$

Conclusion

$$d(X_t^{\mathrm{id}}, X^{\tau}) = \begin{cases} 0 & \text{pba. } \theta(c)^2 \\ 1 & \text{pba. } 1-\theta(c)^2 \end{cases}$$

i.e. $\kappa_c \geq 1 - \mathbb{E}[d(X_t^{\mathrm{id}}, X^{\tau})] = \theta(c)^2$.

Thank You!