Curvature bounds for Markov chains via geodesic convexity: Consequences and examples

Matthias Erbar

joint work (in progress) with C. Henderson, J. Maas, G. Menz, P. Tetali

Universität Bonn

Marne-la-Vallée, October 26–30, 2015

Synthetic notion of lower Ricci bounds

Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) $\mathrm{Ric} \geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

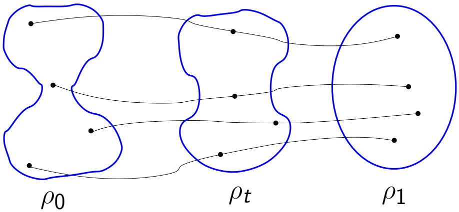
Synthetic notion of lower Ricci bounds

Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) $\mathrm{Ric} \geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

L^2-transport distance:

$$W_2(\rho_0, \rho_1)^2 = \inf \left\{ \int_{M \times M} d(x,y)^2 \mathrm{d}\gamma(x,y) : \gamma \text{ coupling of } \rho_0, \rho_1 \right\}$$

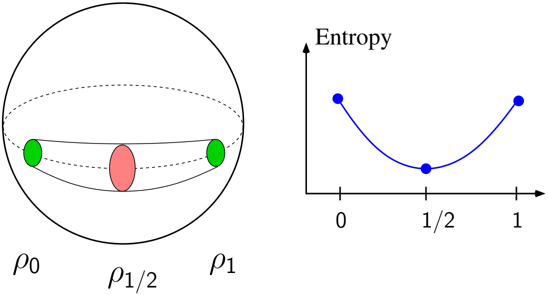


Synthetic notion of lower Ricci bounds

Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) $\mathrm{Ric} \geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

Boltzmann–Shannon entropy $\mathrm{Ent}(\rho) = \int \rho \log \rho \mathrm{dvol}$



Synthetic notion of lower Ricci bounds

Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) Ric $\geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

- (2) makes sense for metric measure spaces
- leads to definition of Ricci curvature lower bounds for mm spaces [Lott-Villani, Sturm]
- many powerful features: stability, functional and geometric inequalities, ...

Synthetic notion of lower Ricci bounds

Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) $\mathrm{Ric} \geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

- (2) makes sense for metric measure spaces
- leads to definition of Ricci curvature lower bounds for mm spaces [Lott-Villani, Sturm]
- many powerful features: stability, functional and geometric inequalities, ...

Problem: unfortunately not applicable for discrete spaces!

Synthetic notion of lower Ricci bounds

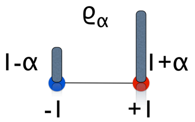
Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) $\mathrm{Ric} \geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

- (2) makes sense for metric measure spaces
- leads to definition of Ricci curvature lower bounds for mm spaces [Lott-Villani, Sturm]
- many powerful features: stability, functional and geometric inequalities, ...

Problem: unfortunately not applicable for discrete spaces!

Example 2-point space



Synthetic notion of lower Ricci bounds

Key observation [Cordero-Erausquin–McCann–Schmuckenschläger, vRenesse–Sturm]

For Riemannian manifold (M, d) are equivalent:
(1) Ric $\geq \kappa$ (2) entropy κ-convex along Wasserstein geodesics

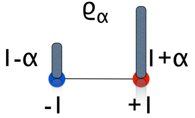
- (2) makes sense for metric measure spaces
- leads to definition of Ricci curvature lower bounds for mm spaces [Lott-Villani, Sturm]
- many powerful features: stability, functional and geometric inequalities, ...

Problem: unfortunately not applicable for discrete spaces!

Example 2-point space

$(\rho^{\alpha(s)})_s$ constant speed geodesic

$|s-t|c = W_2(\rho^{\alpha(s)}, \rho^{\alpha(t)}) = \sqrt{|\alpha(s)-\alpha(t)|}$



Markov chains

Setting

- $\mathcal{X}$ finite set
- $Q : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ transition rates
- L generator of continuous time Markov chain

$$L\psi(x) = \sum_y \big(\psi(y) - \psi(x)\big)\, Q(x,y)$$

- π reversible probability measure on $\mathcal{X}$

$$\forall x, y : Q(x,y)\pi(x) = Q(y,x)\pi(y) \ .$$

Markov chains

Setting

- $\mathcal{X}$ finite set
- $Q : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_+$ transition rates
- L generator of continuous time Markov chain

$$L\psi(x) = \sum_y \big(\psi(y) - \psi(x)\big) Q(x, y)$$

- π reversible probability measure on $\mathcal{X}$

$$\forall x, y : Q(x, y)\pi(x) = Q(y, x)\pi(y) \ .$$

Relative entropy

For $\rho \in \mathcal{P}(\mathcal{X}) := \Big\{ \rho : \mathcal{X} \to \mathbb{R}_+ \ | \quad \sum_{x\in\mathcal{X}} \rho(x)\pi(x) = 1 \Big\}$, define

$$\mathrm{Ent}(\rho) = \textstyle\sum_{x\in\mathcal{X}} \rho(x) \log \rho(x)\pi(x) \ .$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

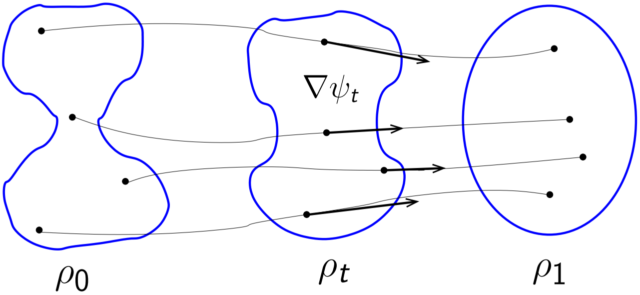
$$W_2(\rho_0, \rho_1)^2 = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla \psi_t(x)|^2 \rho_t(x) \mathrm{d}x \mathrm{d}t \ : \partial_t \rho + \nabla \cdot (\rho \nabla \psi) = 0 \right\} .$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{\rho,\psi}\left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2\,\rho_t(x)\mathrm{d}x\mathrm{d}t \ : \partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0 \right\}.$$



The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0, \rho_1)^2 = \inf_{\rho,\psi} \left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \rho_t(x) \mathrm{d}x\mathrm{d}t \ : \partial_t \rho + \nabla\cdot(\rho\nabla\psi) = 0 \right\} .$$

discrete analogue

$$\mathcal{W}(\rho_0, \rho_1)^2 :=$$

$$\inf_{\rho,\psi} \Bigg\{ \qquad : \qquad \Bigg\} .$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{\rho,\psi}\left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \rho_t(x)\mathrm{d}x\mathrm{d}t \ : \partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0 \right\} .$$

discrete analogue

$$\mathcal{W}(\rho_0,\rho_1)^2 :=$$

$$\inf_{\rho,\psi}\left\{ \int_0^1 \quad \mathrm{d}t \ : \quad \right\} .$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{\rho,\psi}\left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \rho_t(x)\mathrm{d}x\mathrm{d}t \ : \partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0\right\}.$$

discrete analogue

Put $\overline{\nabla}\psi(x,y) = \psi(y) - \psi(x)$, $\mathcal{W}(\rho_0,\rho_1)^2 :=$

$$\inf_{\rho,\psi}\left\{ \int_0^1 \quad |\overline{\nabla}\psi_t(x,y)|^2 \quad \mathrm{d}t \ : \quad \right\}.$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{\rho,\psi}\left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \rho_t(x) \mathrm{d}x\mathrm{d}t \ : \partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0 \right\}.$$

discrete analogue

Put $\overline{\nabla}\psi(x,y) = \psi(y) - \psi(x)$, Fix $\theta : \mathbb{R}_+^2 \to \mathbb{R}_+$. $\mathcal{W}(\rho_0,\rho_1)^2 :=$

$$\inf_{\rho,\psi}\left\{ \int_0^1 \left|\overline{\nabla}\psi_t(x,y)\right|^2 \theta\big(\rho_t(x),\rho_t(y)\big) \mathrm{d}t \ : \right.$$

$$\left. \vphantom{\int} \right\}.$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{\rho,\psi}\left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \rho_t(x)\mathrm{d}x\mathrm{d}t \ : \partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0\right\}.$$

discrete analogue

Put $\overline{\nabla}\psi(x,y) = \psi(y) - \psi(x)$, Fix $\theta : \mathbb{R}_+^2 \to \mathbb{R}_+$. $\mathcal{W}(\rho_0,\rho_1)^2 :=$

$$\inf_{\rho,\psi}\left\{ \frac{1}{2}\int_0^1 \sum_{x,y\in\mathcal{X}} \left|\overline{\nabla}\psi_t(x,y)\right|^2 \theta\big(\rho_t(x),\rho_t(y)\big) Q(x,y)\pi(x)\mathrm{d}t \ : \right.$$

$$\left. \right\}.$$

The discrete transport distance

Recall dynamic characterization of W_2 in $\mathbb{R}^n$

Benamou–Brenier formula in $\mathbb{R}^n$

$$W_2(\rho_0,\rho_1)^2 = \inf_{\rho,\psi}\left\{ \int_0^1 \int_{\mathbb{R}^n} |\nabla\psi_t(x)|^2 \rho_t(x)\mathrm{d}x\mathrm{d}t \ : \partial_t\rho + \nabla\cdot(\rho\nabla\psi) = 0 \right\}.$$

discrete analogue

Put $\overline{\nabla}\psi(x,y) = \psi(y) - \psi(x)$, Fix $\theta : \mathbb{R}_+^2 \to \mathbb{R}_+$. $\mathcal{W}(\rho_0,\rho_1)^2 :=$

$$\inf_{\rho,\psi}\left\{ \frac{1}{2}\int_0^1 \sum_{x,y\in\mathcal{X}} \left|\overline{\nabla}\psi_t(x,y)\right|^2 \theta\big(\rho_t(x),\rho_t(y)\big) Q(x,y)\pi(x)\mathrm{d}t \ : \right.$$

$$\left. \frac{\mathrm{d}}{\mathrm{d}t}\rho_t(x) + \sum_{y\in\mathcal{X}} \theta\big(\rho_t(x),\rho_t(y)\big)\overline{\nabla}\psi_t(x,y)Q(x,y) = 0 \qquad \forall x\in\mathcal{X} \right\}.$$

Properties of the distance $\mathcal{W}$

Choice of θ?

$$\theta(s,t) := \int_0^1 s^{1-\alpha} t^{\alpha} \mathrm{d}\alpha = \frac{s-t}{\log s - \log t} \qquad \text{(logarithmic mean)}$$

Properties of the distance $\mathcal{W}$

Choice of θ?

$$\theta(s,t) := \int_0^1 s^{1-\alpha} t^{\alpha} \mathrm{d}\alpha = \frac{s-t}{\log s - \log t} \qquad \text{(logarithmic mean)}$$

Theorem [Maas '11]

- $\mathcal{W}$ defines geodesic distance on $\mathcal{P}(\mathcal{X})$
- $\mathcal{P}_*(\mathcal{X}) = \{\rho \in \mathcal{P}(\mathcal{X}) \ : \ \rho > 0\}$ is Riem. manifold with distance $\mathcal{W}$
- Semigroup $P_t = e^{tL}$ is the gradient flow of the entropy:

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t \rho = -\operatorname{grad}_{\mathcal{W}} \operatorname{Ent}(P_t \rho) \ .$$

Properties of the distance $\mathcal{W}$

Choice of θ?

$$\theta(s,t) := \int_0^1 s^{1-\alpha} t^{\alpha} \mathrm{d}\alpha = \frac{s-t}{\log s - \log t} \qquad \text{(logarithmic mean)}$$

Theorem [Maas '11]

- $\mathcal{W}$ defines geodesic distance on $\mathcal{P}(\mathcal{X})$
- $\mathcal{P}_*(\mathcal{X}) = \{\rho \in \mathcal{P}(\mathcal{X}) \ : \ \rho > 0\}$ is Riem. manifold with distance $\mathcal{W}$
- Semigroup $P_t = e^{tL}$ is the gradient flow of the entropy:

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t\rho = -\,\mathrm{grad}_{\mathcal{W}} \,\mathrm{Ent}(P_t\rho) \ .$$

discrete analogue of Jordan–Kinderlehrer–Otto '98:

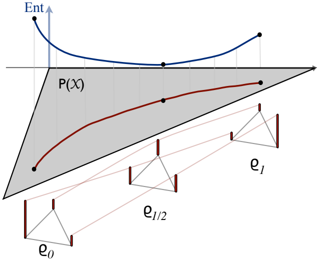
"The heat flow in $\mathbb{R}^d$ is the gradient flow of the entropy w.r.t. W_2"

Entropic curvature bounds for Markov chains

Definition

Markov triple $(\mathcal{X}, Q, \pi)$ has *Ricci bounded below by* $\kappa \in \mathbb{R}$ if the entropy is geodesically κ-convex along geodesics on $(\mathcal{P}(\mathcal{X}), \mathcal{W})$, i.e. for any geodesic $(\rho_t)_{t\in[0,1]}$:

$$\mathrm{Ent}(\rho_t) \le (1-t)\,\mathrm{Ent}(\rho_0) + t\,\mathrm{Ent}(\rho_1) - \frac{\kappa}{2} t(1-t)\mathcal{W}^2(\rho_0, \rho_1)\ .$$

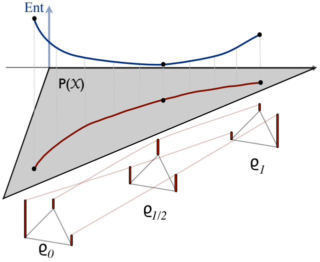


Entropic curvature bounds for Markov chains

Definition

Markov triple $(\mathcal{X}, Q, \pi)$ has *Ricci bounded below by* $\kappa \in \mathbb{R}$ if the entropy is geodesically κ-convex along geodesics on $(\mathcal{P}(\mathcal{X}), \mathcal{W})$, i.e. for any geodesic $(\rho_t)_{t\in[0,1]}$:

$$\mathrm{Ent}(\rho_t) \leq (1-t)\,\mathrm{Ent}(\rho_0) + t\,\mathrm{Ent}(\rho_1) - \frac{\kappa}{2} t(1-t)\mathcal{W}^2(\rho_0, \rho_1)\ .$$

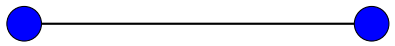


alternative approaches: e.g. Ollivier, Bonciocat–Sturm, Bauer–S.T. Yau et. al

First examples

- **2-point space**

$\mathcal{X} = \{0, 1\}, \quad Q = 1$

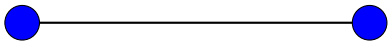


$$\text{Ric} \geq 2$$

First examples

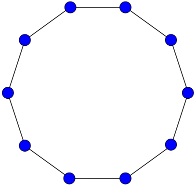
- **2-point space**

$\mathcal{X} = \{0, 1\}, \quad Q = 1$



$$\text{Ric} \geq 2$$

- **Circle**



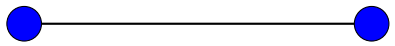
$\mathcal{X} = \{1, \ldots, n\}, \quad Q(i, i \pm 1) = 1$

$$\text{Ric} \geq 0$$

First examples

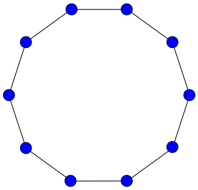
- **2-point space**

$\mathcal{X} = \{0, 1\}, \quad Q = 1$



$$\text{Ric} \geq 2$$

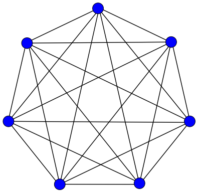
- **Circle**



$\mathcal{X} = \{1, \ldots, n\}, \quad Q(i, i \pm 1) = 1$

$$\text{Ric} \geq 0$$

- **Complete graph**



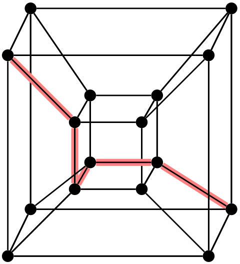
$\mathcal{X} = \{1, \ldots, n\}, \quad Q(i, j) = 1$

$$\text{Ric} \geq \frac{n}{2}$$

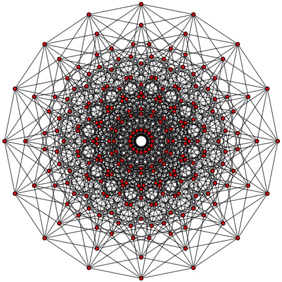
First examples

- **Discrete cube**
  $\mathcal{Q}_n = \{0,1\}^n, \quad Q(x,y) = 1$ for all $x \sim y$

  $\text{Ric} \geq 2$



$\mathcal{Q}_4$



$\mathcal{Q}_8$

Tensorization

For $i = 1, 2$ let $(\mathcal{X}_i, Q_i, \pi_i)$ be Markov triples. Define the product chain $Q_1 \otimes Q_2$ on the product space $\mathcal{X}_1 \times \mathcal{X}_2$ via

$$Q_1 \otimes Q_2\big((x_1, x_2), (y_1, y_2)\big) \;=\; \begin{cases} Q_1(x_1, y_1)\ , & x_2 = y_2\ , \\ Q_2(x_2, y_2)\ , & x_1 = y_1\ , \\ 0\ , & \text{else}\ . \end{cases}$$

$Q_1 \otimes Q_2$ is reversible w.r.t. product measure $\pi_1 \otimes \pi_2$.

Tensorization

For $i = 1, 2$ let $(\mathcal{X}_i, Q_i, \pi_i)$ be Markov triples. Define the product chain $Q_1 \otimes Q_2$ on the product space $\mathcal{X}_1 \times \mathcal{X}_2$ via

$$Q_1 \otimes Q_2\big((x_1, x_2), (y_1, y_2)\big) \; = \; \begin{cases} Q_1(x_1, y_1) \; , & x_2 = y_2 \; , \\ Q_2(x_2, y_2) \; , & x_1 = y_1 \; , \\ 0 \; , & \text{else} \; . \end{cases}$$

$Q_1 \otimes Q_2$ is reversible w.r.t. product measure $\pi_1 \otimes \pi_2$.

Theorem

If $\mathrm{Ric}(\mathcal{X}_i, Q_i, \pi_i) \geq \kappa_i$ for $i = 1, 2$, then

$$\mathrm{Ric}(\mathcal{X}_1 \times \mathcal{X}_2, Q_1 \otimes Q_2, \pi_1 \otimes \pi_2) \geq \min(\kappa_1, \kappa_2) \; .$$

Proving Ricci bounds

Rewrite the Markov chain in *mapping representation*:

- $G \subset \{\delta : \mathcal{X} \to \mathcal{X}\}$ set of allowed moves
- transition rate from x to δx given by $c(x, \delta)$
- Generator rewrites as

$$L\psi(x) = \sum_{\delta \in G} \left[\psi(\delta x) - \psi(x)\right] c(x, \delta)$$

Proving Ricci bounds

Rewrite the Markov chain in *mapping representation*:

- $G \subset \{\delta : \mathcal{X} \to \mathcal{X}\}$ set of allowed moves
- transition rate from x to δx given by $c(x, \delta)$
- Generator rewrites as

$$L\psi(x) = \sum_{\delta \in G} \left[\psi(\delta x) - \psi(x)\right] c(x, \delta)$$

Theorem: If $(\mathcal{X}, Q, \pi)$ has a mapping representation (G, c) satisfying

(1) $\delta \circ \eta = \eta \circ \delta \quad \forall \delta, \eta \in G$ (2) $c(\delta x, \eta) = c(x, \eta) \quad \forall x \in \mathcal{X}, \delta, \eta \in G$,

then $\mathrm{Ric}(Q) \geq 0$.

Proving Ricci bounds

Rewrite the Markov chain in *mapping representation*:

- $G \subset \{\delta : \mathcal{X} \to \mathcal{X}\}$ set of allowed moves
- transition rate from x to δx given by $c(x, \delta)$
- Generator rewrites as

$$L\psi(x) = \sum_{\delta \in G} \left[\psi(\delta x) - \psi(x)\right] c(x, \delta)$$

Theorem: If $(\mathcal{X}, Q, \pi)$ has a mapping representation (G, c) satisfying

(1) $\delta \circ \eta = \eta \circ \delta \quad \forall \delta, \eta \in G$ (2) $c(\delta x, \eta) = c(x, \eta) \quad \forall x \in \mathcal{X}, \delta, \eta \in G$,

then $\mathrm{Ric}(Q) \geq 0$.

Example: Random walks on abelian Cayley graphs

Functional inequalities

Dirichlet form associated to L

$$\mathcal{E}(\varphi, \psi) = \frac{1}{2} \sum \big(\varphi(y) - \varphi(x)\big)\big(\psi(y) - \psi(x)\big) Q(x, y)\pi(x) \ .$$

Along the 'heat semigroup' $P_t = e^{tL}$ we have: $\dfrac{\mathrm{d}}{\mathrm{d}t} \mathrm{Ent}(P_t\rho) = -\mathcal{E}(\rho, \log \rho)$

Functional inequalities

Dirichlet form associated to L

$$\mathcal{E}(\varphi, \psi) = \frac{1}{2} \sum \big(\varphi(y) - \varphi(x)\big)\big(\psi(y) - \psi(x)\big) Q(x, y)\pi(x) \ .$$

Discrete Bakry–Émery theorem

Ric $\geq \lambda > 0$ implies *modified* log-*Sobolev inequality* MLSI(λ):

$$\mathrm{Ent}(\rho) \ \leq \ \frac{1}{2\lambda}\mathcal{E}(\rho, \log \rho) \qquad \forall \rho \in \mathcal{P}_*(\mathcal{X}) \ .$$

Functional inequalities

Dirichlet form associated to L

$$\mathcal{E}(\varphi,\psi) = \frac{1}{2}\sum \big(\varphi(y)-\varphi(x)\big)\big(\psi(y)-\psi(x)\big)\,Q(x,y)\pi(x)\ .$$

Discrete Bakry–Émery theorem

Ric $\geq \lambda > 0$ implies *modified* log-*Sobolev inequality* MLSI(λ):

$$\mathrm{Ent}(\rho) \;\leq\; \frac{1}{2\lambda}\mathcal{E}(\rho,\log\rho) \qquad \forall \rho \in \mathcal{P}_*(\mathcal{X})\ .$$

Discrete Otto–Villani theorem

MLSI(λ) implies the $\mathcal{W}$-*Talagrand inequality* $\mathrm{T}_{\mathcal{W}}(\lambda)$:

$$\mathcal{W}(\rho,\mathbf{1})^2 \;\leq\; \frac{2}{\lambda}\,\mathrm{Ent}(\rho) \qquad \forall \rho \in \mathcal{P}(\mathcal{X})\ .$$

More on functional inequalities

- Both MLSI(λ) and $\mathrm{T}_{\mathcal{W}}(\lambda)$ imply Poincaré inequality P(λ)

$$||\varphi||^2_{L^2(\pi)} \leq \frac{1}{\lambda}\mathcal{E}(\varphi,\varphi) \quad \forall\varphi \text{ s.t. } \sum_x \varphi(x)\pi(x) = 0$$

by linearization

More on functional inequalities

- Both MLSI(λ) and $\mathrm{T}_{\mathcal{W}}(\lambda)$ imply Poincaré inequality P(λ)

$$||\varphi||^2_{L^2(\pi)} \leq \frac{1}{\lambda}\mathcal{E}(\varphi,\varphi) \quad \forall\varphi \text{ s.t. } \sum_x \varphi(x)\pi(x) = 0$$

  by linearization

- $T_{\mathcal{W}}$ implies usual L^1-transport cost inequality and thus exponential concentration for the measure π

More on functional inequalities

- Both MLSI(λ) and T$_{\mathcal{W}}$(λ) imply Poincaré inequality P(λ)

$$||\varphi||^2_{L^2(\pi)} \leq \frac{1}{\lambda}\mathcal{E}(\varphi,\varphi) \quad \forall\varphi \text{ s.t. } \sum_x \varphi(x)\pi(x) = 0$$

  by linearization
- $T_{\mathcal{W}}$ implies usual L^1-transport cost inequality and thus exponential concentration for the measure π
- Trend to equilibrium

$$\begin{aligned}
\mathrm{Ric}(K) \geq \lambda \quad &\Rightarrow \quad \mathcal{W}(P_t\rho, \mathbf{1}) \;\leq\; e^{-\lambda t}\,\mathcal{W}(\rho,\mathbf{1}) \\
MLSI(\lambda) \quad &\Rightarrow \quad \mathcal{H}(P_t\rho) \;\leq\; e^{-2\lambda t}\,\mathcal{H}(\rho) \\
P(\lambda) \quad &\Rightarrow \quad \|P_t\rho - \mathbf{1}\|^2_{L^2(\pi)} \;\leq e^{-\lambda t}\;\|\rho - \mathbf{1}\|^2_{L^2(\pi)}
\end{aligned}$$

Gradient estimates

Bakry–Émery gradient estimate

On Riemannian manifold M with $\mathrm{Ric} \geq \kappa$ we have:

$$|\nabla P_t \psi|^2 \ \leq \ e^{-2\kappa t} P_t \, |\nabla \psi|^2 \quad \forall \psi \ .$$

Gradient estimates

Bakry–Émery gradient estimate

On Riemannian manifold M with $\mathrm{Ric} \geq \kappa$ we have:

$$\int_M |\nabla P_t \psi|^2 \rho \ \leq \ e^{-2\kappa t} \int_M |\nabla \psi|^2 P_t \rho \quad \forall \psi, \rho \, .$$

Gradient estimates

Bakry–Émery gradient estimate

On Riemannian manifold M with $\mathrm{Ric} \geq \kappa$ we have:

$$\int_M |\nabla P_t \psi|^2 \rho \;\leq\; e^{-2\kappa t} \int_M |\nabla \psi|^2 P_t \rho \quad \forall \psi, \rho\,.$$

Theorem

For a Markov triple $(\mathcal{X}, Q, \pi)$, $\mathrm{Ric} \geq \kappa$ is equivalent to the gradient estimate

$$\left|\bar{\nabla} P_t \psi\right|_\rho^2 \;\leq\; e^{-2\kappa t} \left|\bar{\nabla}\psi\right|_{P_t\rho}^2\,.$$

Here we put

$$|\psi|_\rho^2 \;=\; \sum_x \big(\psi(x)\big)^2 \rho(x)\pi(x)\,,$$

$$\left|\bar{\nabla}\psi\right|_\rho^2 \;=\; \frac{1}{2}\sum_{x,y} \big(\psi(y) - \psi(x)\big)^2 \theta\big(\rho(x), \rho(y)\big)\, Q(x,y)\pi(x)\,.$$

Application: Buser inequality

For $A \subset \mathcal{X}$ define the *boundary measure* via

$$|\partial A| \;=\; \sum_{x,y} \mathbf{1}_A(x)\mathbf{1}_{A^c}(y)Q(x,y)\pi(x)\;.$$

Application: Buser inequality

For $A \subset \mathcal{X}$ define the *boundary measure* via

$$|\partial A| \;=\; \sum_{x,y} \mathbf{1}_A(x)\mathbf{1}_{A^c}(y)Q(x,y)\pi(x)\,.$$

Theorem (Buser-Ledoux inequality)

Assume $\mathrm{Ric}(\mathcal{X}, Q, \pi) \geq 0$ and let λ denote the spectral gap of L. Then

$$|\partial A| \;\geq\; \frac{\sqrt{\lambda}}{3\sqrt{Q_*}}\pi(A)\big(1-\pi(A)\big) \quad \forall A \subset \mathcal{X}\,,$$

where

$$Q_* = \min\{Q(x,y) : x, y \ s.t. \ Q(x,y) > 0\}\ .$$

Application: Buser inequality

For $A \subset \mathcal{X}$ define the *boundary measure* via

$$|\partial A| \;=\; \sum_{x,y} \mathbf{1}_A(x)\mathbf{1}_{A^c}(y)Q(x,y)\pi(x)\;.$$

Theorem (Buser-Ledoux inequality)

Assume $\mathrm{Ric}(\mathcal{X},Q,\pi) \geq 0$ and let λ denote the spectral gap of L. Then

$$|\partial A| \;\geq\; \frac{\sqrt{\lambda}}{3\sqrt{Q_*}}\pi(A)\big(1-\pi(A)\big) \quad \forall A \subset \mathcal{X}\;,$$

where

$$Q_* = \min\{Q(x,y) : x,y \text{ } s.t. \text{ } Q(x,y) > 0\}\;.$$

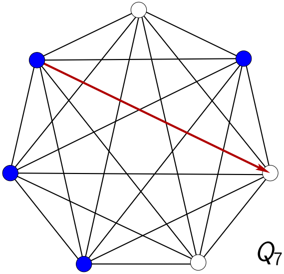
alternative approaches by Klartag–Kozma–Ralli–Tetali and Oveis-Gharan–Trevisan

Ricci bounds for interacting Markov chains

Simple exclusion process

- $1 \leq k \leq n-1$ particles on n sites
- in each step move any particle to any free site at rate $Q_{n,k} = 1/\big(k(n-k)\big)$

Thm: $\mathrm{Ric}(Q_{n,k}) \;\geq\; (n+2)/\big(2k(n-k)\big).$



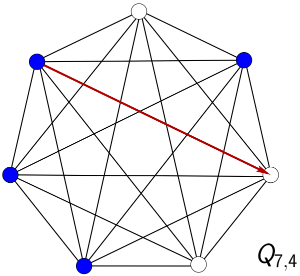
$Q_{7,4}$

Ricci bounds for interacting Markov chains

Simple exclusion process

- $1 \leq k \leq n-1$ particles on n sites
- in each step move any particle to any free site at rate $Q_{n,k} = 1/\big(k(n-k)\big)$

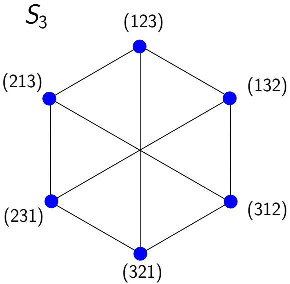
Thm: $\mathrm{Ric}(Q_{n,k}) \ \geq \ (n+2)/\big(2k(n-k)\big).$



Random transposition shuffle

- random walk on symmetric group $\mathcal{S}_n$
- in each step apply random transposition at rate $Q_n = 2/\big(n(n-1)\big)$

Thm: $\mathrm{Ric}(Q_n) \ \geq \ 4/\big(n(n-1)\big).$



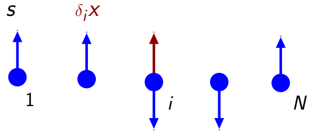
Ricci bounds for interacting Markov chains

Glauber dynamics for general Ising models

To each $i \in 1, \ldots, N$ assign spin ± 1

- State space $\mathcal{X} = \{-1, 1\}^N$
- Energy $H(x) = -\sum_{i,j} \omega_{ij} x_i x_j$ with weigths $\omega_{ij} \in \mathbb{R}$
- Gibbs measure $\pi_\beta(x) = Z^{-1} e^{-\beta H(x)}$ at inverse temperature β
- Glauber dynamics: at each step choose site i at random and flip spin

$$Q_\beta(x, \delta_i x) = \sqrt{\pi_\beta(\delta_i x)/\pi_\beta(x)}$$



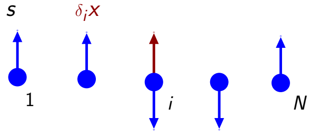
Ricci bounds for interacting Markov chains

Glauber dynamics for general Ising models

To each $i \in 1, \ldots, N$ assign spin ± 1

- State space $\mathcal{X} = \{-1, 1\}^N$
- Energy $H(x) = -\sum_{i,j} \omega_{ij} x_i x_j$ with weigths $\omega_{ij} \in \mathbb{R}$
- Gibbs measure $\pi_\beta(x) = Z^{-1} e^{-\beta H(x)}$ at inverse temperature β
- Glauber dynamics: at each step choose site i at random and flip spin

$$Q_\beta(x, \delta_i x) = \sqrt{\pi_\beta(\delta_i x)/\pi_\beta(x)}$$



Thm: $\mathrm{Ric}(Q_\beta) \geq \kappa(\beta, \omega) > 0$ provided that β is sufficiently small

Thank you for your attention!