

# Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature

M. Erbar and M. Fathi

Matilde Dolfato  
Mentor: Prof. Elia Brué

Università Bocconi  
Visiting Student Initiative - BIDSA

November 22, 2025

We present the paper  
*Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature*  
as a follow up of  
*Ricci curvature of finite markov chains via convexity of the entropy*

We present the paper

*Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature*  
as a follow up of

*Ricci curvature of finite markov chains via convexity of the entropy*

1. The main goal (of the first paper) is to introduce a new geometric structure in the study of Markov semigroups (Markov chains)

We present the paper

*Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature*  
as a follow up of

*Ricci curvature of finite markov chains via convexity of the entropy*

1. The main goal (of the first paper) is to introduce a new geometric structure in the study of Markov semigroups (Markov chains)
2. This is done via an interpretation of the semigroup as a gradient flow of the Shannon entropy
3. Leading to a synthetic notion of Ricci curvature to study contraction properties

We present the paper

*Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature*  
as a follow up of

*Ricci curvature of finite markov chains via convexity of the entropy*

1. The main goal (of the first paper) is to introduce a new geometric structure in the study of Markov semigroups (Markov chains)
2. This is done via an interpretation of the semigroup as a gradient flow of the Shannon entropy
3. Leading to a synthetic notion of Ricci curvature to study contraction properties

**New ingredients (content of the new paper):** functional inequalities and applications to bounds on the mixing time of the chain and to its convergence to equilibrium

## Road from old paper to present

- ▶ Geometric structure on space of Markov chains
- ▶ Synthetic notion of Ricci curvature ( $\text{Ric} \geq 0$ )
- ▶ Analogous version of functional inequalities for Markov chains
- ▶ Bounds on mixing time and rate of convergence to equilibrium  
(among others)

# Overview of the talk

## Introduction

Main result: convergence to equilibrium of the zero-range process

The zero-range process

Bound on mixing time and interpretation

## Synthetic geometric structure

Setup

Distance  $\mathcal{W}$

Ricci curvature for discrete processes

## Modified log-Sobolev inequality

## From MLSI to convergence to equilibrium

# The zero-range process I

Consider  $K$  interacting particles on the complete graph with  $L$  sites.

The state space is  $\mathcal{X}_{K,L} = \{\eta \in \mathbb{N}^L : \sum_i \eta_i = K\}$ .

The zero-range process with **constant rates** can be described as:

1. Choose a site  $i$  uniformly at random
  - 1.1 If  $\eta_i = 0$ , do nothing
  - 1.2 Else, choose another site  $j$  uniformly at random and move 1 particle from  $i$  to  $j$

$\eta^{i,j}$  denotes the new configuration after such a move.

$$Q_{K,L}(\eta, \theta) = \begin{cases} \frac{1}{L} & \theta = \eta^{i,j} \text{ for some } i, j \\ 0 & \text{else} \end{cases}$$

Invariant measure is the uniform measure  $\pi_{K,L}$ .

## The zero-range process II

Good model for:

- ▶ traffic flow
- ▶ population dynamics
- ▶ particle physics
- ▶ ... any system in which deciding to move depends on the surroundings

EB: se trovi il nome specifico di modelli in traffico flow, population design, etc. aggiungili

## Main result and interpretation

### Convergence to equilibrium

The constant-rate zero range process with  $K$  particles and  $L$  sites has mixing time bounded by

$$\tau_{\text{mix}}(\varepsilon) \leq KL \log L \left( \frac{1}{8} - \frac{\log \varepsilon}{c} \right) \quad (1)$$

for some universal constant  $c$ .

## Main result and interpretation

### Convergence to equilibrium

The constant-rate zero range process with  $K$  particles and  $L$  sites has mixing time bounded by

$$\tau_{\text{mix}}(\varepsilon) \leq KL \log L \left( \frac{1}{8} - \frac{\log \varepsilon}{c} \right) \quad (1)$$

for some universal constant  $c$ .

\* This is exponential convergence to equilibrium! \*

# Overview of the talk

Introduction

Main result: convergence to equilibrium of the zero-range process

The zero-range process

Bound on mixing time and interpretation

Synthetic geometric structure

Setup

Distance  $\mathcal{W}$

Ricci curvature for discrete processes

Modified log-Sobolev inequality

From MLSI to convergence to equilibrium

## Introduction

Main result: convergence to equilibrium of the zero-range process

The zero-range process

Bound on mixing time and interpretation

## Synthetic geometric structure

Setup

Distance  $\mathcal{W}$

Ricci curvature for discrete processes

Modified log-Sobolev inequality

From MLSI to convergence to equilibrium

## Setup

Finite space of states  $\mathcal{X}$ , irreducible, reversible Markov kernel  $Q$  and stationary measure  $\pi$

$$Q(x, y)\pi(x) = Q(y, x)\pi(y)$$

Operator  $L$  acting on functions  $\psi : \mathcal{X} \rightarrow \mathbb{R}$  is

$$L\psi(x) = \sum_{y \in \mathcal{X}} (\psi(y) - \psi(x))Q(x, y)$$

is the generator of a continuous time Markov chain.

Space of probability densities

$$\mathcal{P}(\mathcal{X}) := \left\{ \rho : \mathcal{X} \rightarrow \mathbb{R}_+ : \sum_{x \in \mathcal{X}} \rho(x) \pi(x) = 1 \right\}$$

## Distance $\mathcal{W}$

We define a discrete transport distance, by analogy with the Wasserstein distance. For  $\rho_0, \rho_1 \in \mathcal{P}(\mathcal{X})$ ,

$$\mathcal{W}(\rho_0, \rho_1)^2 := \inf_{\rho, \psi} \frac{1}{2} \int_0^1 \sum_{x, y \in \mathcal{X}} (\psi_t(x) - \psi_t(y))^2 \hat{\rho}_t(x, y) Q(x, y) \pi(x) dt$$

where the infimum runs over all sufficiently regular curves satisfying a continuity equation

$$\begin{cases} \frac{d}{dt} \rho_t(x) + \sum_{y \in \mathcal{X}} (\psi_t(y) - \psi_t(x)) \hat{\rho}_t(x, y) Q(x, y) = 0, & \forall x \in \mathcal{X}, \\ \rho|_{t=0} = \rho_0, \quad \rho|_{t=1} = \rho_1. \end{cases}$$

and  $\hat{\rho}(x, y) := \theta(\rho(x), \rho(y))$  which is the logarithmic mean.

# Markov semigroup as gradient flow

$$\mathcal{W}(\rho_0, \rho_1)^2 := \inf_{\rho, \psi} \frac{1}{2} \int_0^1 \sum_{x, y \in X} (\psi_t(x) - \psi_t(y))^2 \hat{\rho}_t(x, y) Q(x, y) \pi(x) dt$$

$$\text{s.t. } \frac{d}{dt} \rho_t(x) + \nabla \cdot (\hat{\rho}(x, y) \nabla \psi(x, y)) = 0$$

Endowing  $\mathcal{X}$  with this distance, we can interpret the Markov semigroup  $P_t = e^{tL}$  as a gradient flow of Shannon's entropy

$$\mathcal{H}(\rho) := \sum_{x \in \mathcal{X}} \rho(x) \log \rho(x) \pi(x)$$

# Ricci for discrete processes

## Entropic Ricci curvature for discrete space

$(\mathcal{X}, Q, \pi)$  has entropic Ricci curvature bounded from below by  $\kappa \in \mathbb{R}$  if for any geodesic  $\{\rho_t\}_{t \in [0,1]}$  on  $(\mathcal{P}(\mathcal{X}), \mathcal{W})$  we have

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

In this case, we write  $\text{Ric}(\mathcal{X}, Q, \pi) \geq \kappa$ .

## Ricci for the zero-range process (in brief)

In another contribution, M. Fathi and J. Maas provide explicit ways of computing bounds on Ricci for discrete processes

They show that for the ZRP with *increasing* rates, Ricci is positive

The ZRP with *constant* rates has  $\text{Ric} \geq 0$

# Overview of the talk

## Introduction

Main result: convergence to equilibrium of the zero-range process

The zero-range process

Bound on mixing time and interpretation

## Synthetic geometric structure

Setup

Distance  $\mathcal{W}$

Ricci curvature for discrete processes

## Modified log-Sobolev inequality

From MLSI to convergence to equilibrium

# The Modified log-Sobolev inequality

## Modified logarithmic Sobolev inequality

It is a **bound on the entropy** in terms of a norm of the gradient of our observable:

$$\mathcal{H}(\rho) \leq \frac{1}{2\lambda} \mathcal{I}(\rho) \quad (\text{MLSI}(\lambda))$$

where  $\mathcal{I}$  is the Fisher information

$$\mathcal{I}(\rho) := \int \rho |\nabla \log \rho|^2 d\pi = \|\nabla \log \rho\|_{\rho}^2$$

# The Modified log-Sobolev inequality

## Modified logarithmic Sobolev inequality

It is a **bound on the entropy** in terms of a norm of the gradient of our observable:

$$\mathcal{H}(\rho) \leq \frac{1}{2\lambda} \mathcal{I}(\rho) \quad (\text{MLSI}(\lambda))$$

where  $\mathcal{I}$  is the Fisher information

$$\mathcal{I}(\rho) := \int \rho |\nabla \log \rho|^2 d\pi = \|\nabla \log \rho\|_{\rho}^2$$

Many implications on convergence property, spectrum of the Laplacian and hypercontractivity!

- \* Our focus: how MLSI allows to bound the mixing time of the chain via entropy decay \*

# Overview of the talk

Introduction

Main result: convergence to equilibrium of the zero-range process

The zero-range process

Bound on mixing time and interpretation

Synthetic geometric structure

Setup

Distance  $\mathcal{W}$

Ricci curvature for discrete processes

Modified log-Sobolev inequality

From MLSI to convergence to equilibrium

## From MLSI to entropy decay

$$\mathcal{H}(\rho) \leq \frac{1}{2\lambda} \mathcal{I}(\rho) \quad (\text{MLSI}(\lambda))$$

When  $P_t$  is the heat semigroup,

$$\frac{d}{dt} \mathcal{H}(P_t \rho) = -\mathcal{I}(P_t \rho)$$

Hence

$$\mathcal{H}(P_t \rho) \leq \frac{1}{2\lambda} \frac{d}{dt} \mathcal{H}(P_t \rho) \quad (\text{MLSI})$$

Using Gronwall's inequality, this yields an estimate on entropy decay

$$\mathcal{H}(P_t \rho) \leq e^{-2\lambda t} \mathcal{H}(\rho)$$

that quantifies convergence to equilibrium.

## From entropy decay to bound on mixing time

We look at convergence to equilibrium in terms of the **mixing time** of the Markov chain.

Total variation mixing time of a Markov chain

For  $\varepsilon > 0$ ,

$$\tau_{\text{mix}}(\varepsilon) := \inf \left\{ t > 0 : \|P_t^* \delta_x - \pi\|_{TV} < \varepsilon \quad \forall x \in \mathcal{X} \right\}$$

Since

$$\|\nu - \pi\|_{TV}^2 \leq \frac{1}{2} \mathcal{H}(\nu) \quad (\text{Pinsker inequality})$$

we can use MLSI to bound  $\tau_{\text{mix}}$ .

Convergence to equilibrium

Assume  $\text{Ric}(\mathcal{X}, Q, \pi) \geq 0$  and the diameter of  $(\mathcal{X}, d_{\mathcal{W}})$  is  $\leq D$ . If a MLSI with constant  $\lambda = \frac{c}{D^2}$  holds, then

$$\tau_{\text{mix}}(\varepsilon) \leq D^2(1/8 - c \log \varepsilon)$$

for some universal constant  $c$

# Modified Log-Sobolev inequality for the zero-range process

## Diameter bound

There is a constant  $c > 0$  such that the diameter of  $(\mathcal{X}_{K,L}, d_{\mathcal{W}})$  is bounded by  $cK\sqrt{L \log L}$

So, for the zero range process with constant uniform rates, a MLSI with constant  $\frac{c}{K^2 L \log L}$  holds.

## Convergence of the ZRP

The total variation mixing time of the zero range process on  $\mathcal{X}_{K,L}$  with constant uniform rates is

$$\tau_{\text{mix}}(\varepsilon) \leq KL \log L \left( \frac{1}{8} - \frac{\log \varepsilon}{c} \right)$$

for some universal constant  $c$ .

## Other examples I

lalaland

# Additional formulas I

## Distance $d_{\mathcal{W}}$ on $\mathcal{X}$

The distance  $\mathcal{W}$  on  $\mathcal{P}(\mathcal{X})$  induces a distance on  $\mathcal{X}$  by restricting to Dirac masses, i.e. for  $x, y \in \mathcal{X}$

$$d_{\mathcal{W}}(x, y) := \mathcal{W}(\delta_x, \delta_y)$$

## Induced Riemannian structure on $\mathcal{X}$

$$(\Psi, \Phi)_\rho := \frac{1}{2} \sum_{x,y} \Psi(x, y) \Phi(x, y) \hat{\rho}(x, y) Q(x, y) \pi(x)$$

## Additional formulas II

### Discrete gradient and divergence

$$\nabla\psi(x, y) = \psi(y) - \psi(x)$$

$$\begin{aligned}\nabla \cdot (\nabla\psi)(x) &= \frac{1}{2} \sum_y (\nabla\psi(x, y) - \nabla\psi(y, x)) Q(x, y) \\ &= \frac{1}{2} \sum_y (\psi(y) - \psi(x) - \psi(x) + \psi(y)) Q(x, y) \\ &= \sum_y (\psi(y) - \psi(x)) Q(x, y)\end{aligned}$$

## Additional formulas III

Heat flow as gradient flow of entropy

Heat equation  $\rho'_t = \Delta \rho_t = \nabla \cdot \nabla \rho_t$

Continuity equation  $\rho'_t + \nabla \cdot (\hat{\rho} \nabla \psi) = 0$

The heat equation can be re-written as a continuity equation if

$$\nabla \psi_t = -\frac{\nabla \rho_t}{\hat{\rho}_t}$$

The gradient of the entropy is  $\text{grad}_{\mathcal{W}} \mathcal{H}(\rho_t) = \nabla \log \rho_t$

Hence, we get that the heat flow is the gradient flow of the entropy if

$$\nabla \psi_t = -\frac{\nabla \rho_t}{\hat{\rho}_t} = -\nabla \log \rho_t;$$

which gives

$$\frac{\nabla \rho_t}{\hat{\rho}_t} = \nabla \log \rho_t \tag{2}$$

i.e. precisely that  $\hat{\rho}_t$  is the logarithmic mean.