

Curvature and transport inequalities for Markov chains in discrete spaces

MAX FATHI¹ and YAN SHU²

¹*Department of Mathematics, UC Berkeley, Berkeley, CA, USA. E-mail: maxf@berkeley.edu*

²*Université Paris-Ouest Nanterre, France. E-mail: yshu@u-paris10.fr*

We study various transport-information inequalities under three different notions of Ricci curvature in the discrete setting: the curvature-dimension condition of Bakry and Émery (In *Séminaire de Probabilités, XIX, 1983/84* (1985) 177–206 Springer), the exponential curvature-dimension condition of Bauer *et al.* (*Li-Yau Inequality on Graphs* (2013)) and the coarse Ricci curvature of Ollivier (*J. Funct. Anal.* **256** (2009) 810–864). We prove that under a curvature-dimension condition or coarse Ricci curvature condition, an L_1 transport-information inequality holds; while under an exponential curvature-dimension condition, some weak-transport information inequalities hold. As an application, we establish a Bonnet–Myers theorem under the curvature-dimension condition $CD(\kappa, \infty)$ of Bakry and Émery (In *Séminaire de Probabilités, XIX, 1983/84* (1985) 177–206 Springer).

Keywords: curvature; discrete spaces; functional inequalities; Markov chains; optimal transport

1. Introduction

In the analysis of the geometry of Riemannian manifolds, Ricci curvature plays an important role. In particular, Ricci curvature lower bounds immediately yield powerful functional inequalities, such as the logarithmic Sobolev inequality, which in turn implies transport-entropy and transport-information inequalities. Each of these inequalities has its own interest and has various applications, such as concentration bounds and estimates on the speed of convergence to equilibrium for Markov chains. We refer the reader to [4,20,45] for more about the links between curvature and functional inequalities, and to [2,27] for applications of functional inequalities.

However, when the space we consider is a graph, those theories are not as clear as in the continuous settings. The first question one would want to answer is how to define Ricci curvature lower bounds in discrete settings. The natural approach would be to define it as a discrete analogue of a definition valid in the continuous setting. There are several equivalent definitions one can try to use (see [1] for those definitions in the continuous settings and for the equivalences between them). However, in discrete spaces, we lose the chain rule, and these definitions are no longer equivalent.

Several notions of curvature have been proposed in the last few years. Here we shall consider three of them: the curvature-dimension condition of [3], the exponential curvature-dimension condition of [5] and the coarse Ricci curvature of [37]. Other notions that have been developed (and which we shall not discuss further here) include the entropic Ricci curvature defined in [14] and [34], which is based on the Lott–Sturm–Villani definition of curvature [30,43], geodesic convexity along interpolations in [17] and [28], rough curvature bounds in [8]. It is still an open

problem to compare these various notions of curvature. We refer readers to the forthcoming survey [11] for a more general introduction.

The aim of this work is to obtain functional inequalities under the above three notions of curvature conditions and give some applications.

Let us begin with setting the framework of Markov chains on discrete spaces:

Markov chain on graphs and curvature condition

Let \mathcal{X} be a finite (or countably infinite) discrete space and K be an irreducible Markov kernel on \mathcal{X} . Assume that for any $x \in \mathcal{X}$, we have

$$\sum_y K(x, y) = 1. \tag{1}$$

This condition is a normalization of the time scale, enforcing that jump attempts occur at rate 1. We also define $J(x) := 1 - K(x, x)$ and $J := \sup_{x \in \mathcal{X}} J(x)$. J is a measure of the laziness of the chain, estimating how often jump attempts end with the particle not moving. Since we assume the kernel is irreducible, $0 < J \leq 1$.

We shall always assume there exists a reversible invariant probability measure π , satisfying the detailed balance relation

$$K(x, y)\pi(x) = K(y, x)\pi(y) \quad \forall x, y \in \mathcal{X}.$$

We denote by L the generator of the continuous-time Markov chain associated to the kernel K , which is given by

$$Lf(x) = \sum_y (f(y) - f(x))K(x, y).$$

Let $P_t = e^{tL}$ be the associated semigroup, acting on functions, and P_t^* its adjoint, acting on measures. We also define the Γ operator, given by

$$\Gamma(f, g)(x) := \frac{1}{2} \sum_y (f(y) - f(x))(g(y) - g(x))K(x, y)$$

and write $\Gamma(f) := \Gamma(f, f)$.

With this Γ operator, we are able to introduce the Bakry–Émery curvature condition $\text{CD}(\kappa, \infty)$:

Definition 1.1. We define the iterated Γ operator Γ_2 as

$$\Gamma_2(f) = \frac{1}{2}L\Gamma(f) - \Gamma(f, Lf).$$

We say that the curvature condition $\text{CD}(\kappa, \infty)$ is satisfied if, for all functions f , we have

$$\Gamma_2(f) \geq \kappa\Gamma(f).$$

Since we shall deal with three different types of Ricci curvature lower bounds, in order to avoid confusions, we always denote κ for the Bakry–Émery curvature condition, κ_e for the exponential curvature dimension condition and κ_c for a lower bound on the Coarse Ricci curvature. Throughout the paper, $\kappa, \kappa_e, \kappa_c$ will always be positive numbers.

The Γ_2 operator and the curvature condition were first introduced in [3], and used to prove functional inequalities, such as logarithmic Sobolev inequalities and Poincaré inequalities, for measures on Riemannian spaces satisfying $\text{CD}(\kappa, \infty)$ for $\kappa > 0$. In the Riemannian setting, the Γ_2 operator involves the Ricci tensor of the manifold, and the condition $\text{CD}(\kappa, \infty)$ is equivalent to asking for lower bounds on the Ricci curvature, and more generally to the Lott–Sturm–Villani definition of lower bounds on Ricci curvature (see [30] and [43] for the definition, and [1] for the equivalence between the two notions). Hence, $\text{CD}(\kappa, \infty)$ can be used as a *definition* of lower bounds on the Ricci curvature for nonsmooth spaces, and even discrete spaces. This was the starting point of a very fruitful direction of research on the links between curvature and functional inequalities. In most cases, the focus was on the continuous setting, and the operator L was assumed to be a diffusion operator.

In the discrete setting, this curvature condition was first studied in [41] and then in [5,6,10,29]. It has also been used in [26], where a discrete version of Buser’s inequality was obtained, as well as curvature bounds for various graphs, such as abelian Cayley graphs and slices of the hypercube. Note that most of these works are set in the framework of graphs rather than Markov chains, which generally makes our definitions and theirs differ by a normalization constant, since we enforce the condition (1).

In the discrete setting, obtaining a lower bound on the Bakry–Émery curvature is somewhat difficult for concrete models. Klartag *et al.* [26] studies several examples of Markov chains for which Bakry–Émery curvature bounds can be proven, such as the random transposition model and the Bernoulli–Laplace process. Such curvature bounds for birth and death processes have also been studied by [22]. The situation is better understood in the continuous setting, where it includes the important case of Langevin samples with uniformly convex potentials.

As we have mentioned, the main differences between the continuous and discrete settings is the when the operator L is not a diffusion operator, we lose the chain rule. This leads to additional difficulties, and some results, such as certain forms of the logarithmic Sobolev inequality, do not seem to hold anymore. On the other hand, one of the main difficulties in the continuous setting is to exhibit an algebra of smooth functions satisfying certain conditions, while this property immediately holds in the discrete setting.

The key chain rule used in the continuous setting is the identity

$$L(\Phi(f)) = \Phi'(f)Lf + \Phi''(f)\Gamma(f)$$

which characterizes diffusion operators in the continuous setting, and does not hold in discrete settings. However, a key observation of [5] is that when $\Phi(x) = \sqrt{x}$, the identity

$$2\sqrt{f}L\sqrt{f} = Lf - 2\Gamma(\sqrt{f})$$

holds, even in the discrete setting. This observation motivated the introduction of a modified version of the curvature-dimension condition, designed to exploit this identity.

Definition 1.2. We define the modified Γ_2 operator $\tilde{\Gamma}_2$ as

$$\tilde{\Gamma}_2(f, f) := \Gamma_2(f) - \Gamma\left(f, \frac{\Gamma(f)}{f}\right).$$

We say that the exponential curvature condition $\text{CDE}'(\kappa_e, \infty)$ is satisfied if, for all nonnegative functions f and all $x \in \mathcal{X}$, we have

$$\tilde{\Gamma}_2(f)(x) \geq \kappa_e \Gamma(f)(x).$$

Remark 1.1. We use the notation $\text{CDE}'(\kappa_e, \infty)$ to agree with the notations of [5], where they also consider the case when the condition is only satisfied at points x where $Lf(x) < 0$.

In [35], it is shown that $\text{CDE}'(\kappa_e, \infty)$ implies $\text{CD}(\kappa, \infty)$ with $\kappa = \kappa_e$. When the operator L is a diffusion, the conditions $\text{CDE}'(\kappa_e, \infty)$ and $\text{CD}(\kappa, \infty)$ are equivalent.

Under this notion of curvature, [5] prove various Li-Yau inequalities on graphs, and then deduce heat kernel estimates and a Buser inequality for graphs. In [6], it was shown that the $\text{CDE}'(\kappa_e, \infty)$ condition tensorizes, and that the associated heat kernel satisfies some Gaussian bounds.

At the time of writing, we do not know many examples for which lower bounds on the modified Bakry-Émery curvature have been computed. In [5], it is shown that Abelian Cayley graphs (and more generally the so-called Ricci-flat graphs) satisfy $\text{CDE}'(0, \infty)$, and we show in Section 6 that the N -dimensional discrete hypercube satisfies $\text{CDE}'(1/N, \infty)$. It would be interesting to study this curvature condition for more examples. A starting point could be the examples studied in [26] for which the weaker $\text{CD}(\kappa, \infty)$ condition has been established.

The third notion of curvature we shall now introduce is the coarse Ricci curvature. In order to define it, we first need to introduce Wasserstein distances.

Let d be a distance on \mathcal{X} . The L^p Wasserstein distance is defined as the following.

Definition 1.3 (L^p -Wasserstein distances). Let $p \geq 1$. The L^p -Wasserstein distance W_p between two probability measures μ and ν on a metric space (\mathcal{X}, d) is defined as

$$W_p(\mu, \nu) := \left(\inf_{\pi} \int d(x, y)^p \pi(dx dy) \right)^{1/p},$$

where the infimum runs over all couplings π of μ and ν .

Finally, we recall the definition of coarse Ricci curvature, which has been introduced in [37] for discrete-time Markov chains. Since we shall work in continuous time, we shall give the appropriate variant, introduced in [24]. Previous works considering contraction rates in transport distance include [12,36,40]. Applications to error estimates for Markov Chain Monte Carlo were studied in [25]. The continuous-time version we use here was introduced in [24]. The particular case of curvature on graphs has been studied in [23].

Definition 1.4 (Coarse Ricci curvature). The coarse Ricci curvature of the Markov chain is said to be bounded from below by κ_c if, for all probability measures μ and ν and any time $t \geq 0$, we have

$$W_1(P_t^* \mu, P_t^* \nu) \leq \exp(-\kappa_c t) W_1(\mu, \nu),$$

that is, if it is a contraction in W_1 distance, with rate κ_c .

Note that unlike the $CD(\kappa, \infty)$ condition, this property does not only depend on the Markov chain, but also on the choice of the distance d .

In this work, as a distance on \mathcal{X} we shall use the *graph distance* associated to the Markov kernel. If we consider \mathcal{X} as the set of vertices of a graph, with edges between all pairs of vertices (x, y) such that $K(x, y) > 0$, d shall be the usual graph distance. More formally, it is defined as

$$d(x, y) := \inf\{n \in \mathbb{N}; \exists x_0, \dots, x_n | x_0 = x, x_n = y, K(x_i, x_{i+1}) > 0 \forall 0 \leq i \leq n - 1\}.$$

Because of the relationship between coarse Ricci curvature and path coupling [9] (a notion commonly used for the study of fast mixing of Markov chains), there are many known examples of Markov chain having positive coarse Ricci curvature. Many of them are surveyed in [38], but we would like to mention a few with concrete interest, such as birth and death chains, Glauber dynamics on the Ising model at high temperature and random walks on manifolds with positive Ricci curvature. See also [31] for applications to concrete models in computer science and statistical physics.

At the time of writing, it is not known whether there is a direct relationship between Bakry–Émery curvature and coarse Ricci curvature in the discrete setting.

Functional inequalities

Now we turn to functional inequalities on graphs.

Definition 1.5 (Fisher information). Let f be a nonnegative function defined on \mathcal{X} . Define the Fisher information I_π of f with respect to π as

$$\mathcal{I}_\pi(f) := 4 \int \Gamma(\sqrt{f}) d\pi = 2 \sum_{x \in cX} \sum_{y \in \mathcal{X}} (\sqrt{f(y)} - \sqrt{f(x)})^2 K(x, y) \pi(x).$$

The factor 4 in this definition comes from the analogy with the continuous setting, where

$$4 \int |\nabla \sqrt{f}|^2 d\pi = \int |\nabla \log f|^2 f d\mu.$$

In the continuous setting, the Fisher information can be written as $\int \nabla \log f \cdot \nabla f d\pi$, so we can define a modified Fisher information as

$$\tilde{\mathcal{I}}_\pi(f) := \int \Gamma(f, \log f) d\pi, \tag{2}$$

which corresponds to the entropy production functional of the Markov chain.

There is a third way to rewrite the Fisher information for the continuous settings as $\int \frac{|\nabla f|^2}{f} d\mu$, and one can also define another modified Fisher information as

$$\bar{\mathcal{I}}_\pi(f) := \int \frac{\Gamma(f)}{f} d\pi.$$

Of course, there are many other ways to rewrite the Fisher information in the continuous setting, each leading to a different definition in the discrete setting. We only stated here the three versions we shall use in this work.

In the discrete setting, $\mathcal{I}_\pi(f)$, $\tilde{\mathcal{I}}_\pi(f)$ and $\bar{\mathcal{I}}_\pi(f)$ are not equal in general. It is easy to see that $\mathcal{I}_\pi(f) \leq \tilde{\mathcal{I}}_\pi(f)$ and $\mathcal{I}_\pi(f) \leq \bar{\mathcal{I}}_\pi(f)$. If f is the density function of a probability measure ν with respect to π , since $(\sqrt{f(y)} - \sqrt{f(x)})^2 \leq f(x) + f(y)$, and since π is reversible, one can deduce that

$$\mathcal{I}_\pi(f) \leq 2 \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{X}} (f(x) + f(y)) K(x, y) \pi(x) \leq 4J.$$

Here we can see that in discrete settings, the Fisher information is in fact bounded from above, which is not true in continuous settings.

Let us recall the definition of the relative entropy Ent_π as well:

Definition 1.6 (Relative entropy). Assuming that f is a nonnegative function on \mathcal{X} , we define the relative entropy f with respect to π as

$$\text{Ent}_\pi(f) := \sum_{\mathcal{X}} f(x) \log f(x) \pi(x) - \sum_{\mathcal{X}} f(x) \pi(x) \log \left(\sum_{\mathcal{X}} f(x) \pi(x) \right).$$

Note that when f is a probability density with respect to π , the second term takes value 0.

Definition 1.7. Let π be a probability measure on \mathcal{X} and $p \geq 1$. We say that π satisfies:

(i) the logarithmic Sobolev inequality with constant C , if for all nonnegative functions f , we have

$$\text{Ent}_\pi(f) \leq \frac{1}{2C} \mathcal{I}_\pi(f) \quad (\text{LSI}(C));$$

(ii) the modified logarithmic Sobolev inequality with constant C , which we shall write $\text{mLSI}(C)$, if for all nonnegative functions f , we have

$$\text{Ent}_\pi(f) \leq \frac{1}{2C} \int \Gamma(f, \log f) d\pi \quad (\text{mLSI}(C));$$

(iii) the transport-entropy inequality $T_p(C)$ if for all probability measures $\nu = f\pi$, we have

$$W_p(\nu, \mu)^2 \leq \frac{2}{C} \text{Ent}_\pi(f) \quad (T_p(C));$$

(iv) the transport-information inequality $T_p I(C)$ if for all probability measures $\nu = f\pi$, we have

$$W_p(\nu, \mu)^2 \leq \frac{1}{C^2} \mathcal{I}_\pi(f) \quad (T_p I(C)).$$

In the continuous setting, mLSI and LSI are the same inequality, but in the discrete setting they correspond to distinct properties of the Markov chain, namely hypercontractivity for LSI and exponential convergence to equilibrium in relative entropy for mLSI. In general, LSI implies mLSI, but the converse is *not* true. We refer to [7] for more on the difference between the two inequalities.

In the discrete setting, when $p = 1$, the following relations between the inequalities hold, in the same way as in the continuous setting:

$$\text{LSI}(C) \Rightarrow T_1 I(C) \Rightarrow T_1(C).$$

The T_1 inequality is equivalent to Gaussian concentration for π (see, for example, [27] and the next section), but not dimension-free concentration, and is therefore strictly weaker than T_2 . When $p = 1$, the transport-information inequality is equivalent to Gaussian concentration for the occupation measure of the Markov chain (see [21]), and is therefore useful to get a priori bounds on the statistical error for Markov Chain Monte Carlo estimation of averages. Gaussian concentration bounds typically take the form

$$\mathbb{P}^\pi (f \geq \mathbb{E}[f] + r) \leq M e^{-cr^2}$$

for any 1-Lipschitz function f and any $r \geq 0$. Hence, such inequalities allow to bound probabilities of observing significant deviations for Lipschitz observables. Such results can sometimes be obtained for non-Lipschitz observables via function inequalities using extra assumptions, see, for example, [19].

The transport-information inequality implies concentration bounds for the occupation measure of a Markov chain. It was shown in [21] to be equivalent to

$$\mathbb{P}^\nu \left[\frac{1}{t} \int_0^t g(X_t) dt \geq \pi(g) + r \right] \leq \left\| \frac{d\nu}{d\pi} \right\|_2 \exp(-tr^2/2C)$$

for any initial data ν , any 1-Lipschitz observable g and any $r > 0$. Other consequences (bounds on Feynmann–Kac semigroups, large deviations) have also been obtained in [21].

One of the most interesting cases in the continuous setting is the transport-entropy inequality when $p = 2$, which is also called the Talagrand inequality and was introduced in [44]. It is equivalent to dimension-free Gaussian concentration for π [15].

One can show the following relationships:

$$\text{CD}(\kappa, \infty) \Rightarrow \text{LSI}(\kappa) \Rightarrow T_2 I(\kappa) \Rightarrow T_2(\kappa).$$

We refer to [39] and [20] for the proofs of these implications. With those inequalities in hand, one can prove some dimension-free concentration results on a metric-measure space (see [15]).

However, the results for $p = 2$ fail to be true in discrete settings, as we will see in next section. When \mathcal{X} is a graph, π never satisfies T_2 , unless it is a Dirac measure (see for example, [16], or Section 2). To recover a discrete version of T_2 , we therefore have to redefine the transport cost. Erbar and Maas recovered some of those functional inequality results with the notion of entropic Ricci curvature on graphs, and we refer the reader to [14,32] for more details. Another way to deal with it is to take the weak transport cost introduced in [33]:

Definition 1.8. Let (\mathcal{X}, d) be a polish space and μ, ν two probabilities measures on \mathcal{X} , define

$$\tilde{\mathcal{T}}_2(\nu|\mu) := \inf_{\pi \in \Pi(\mu, \nu)} \left\{ \int \left(\int d(x, y) p_x(dy) \right)^2 \mu(dx) \right\}.$$

Where $\Pi(\mu, \nu)$ is the set of all couplings π whose first marginal is μ and second marginal is ν , p_x is the probability kernel such that $\pi(dxdy) = p_x(dy)\mu(dx)$. Using probabilistic notations, one has

$$\tilde{\mathcal{T}}_2(\nu|\mu) = \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}((\mathbb{E}(d(X, Y)|X))^2).$$

Note that the weak transport cost could be also seen as a weak Wasserstein-like distance. In order to agree with the notations of Wasserstein distance, we note $\tilde{W}_2(\nu|\mu)^2 := \tilde{\mathcal{T}}_2(\nu|\mu)$. However, it is *not* a distance, since it is not symmetric. Note that by Jensen’s inequality, both $\tilde{W}_2(\nu|\mu)$ and $\tilde{W}_2(\mu, \nu)$ are larger than $W_1(\mu, \nu)$.

Definition 1.9. Adapting the settings of Definition 1.7, we say that π satisfies:

(v) the weak transport-entropy inequality $\tilde{T}_2^+(C)$ if for all probability measures $\nu = f\pi$, we have

$$\tilde{W}_2(f\pi|\pi)^2 \leq \frac{2}{C} \text{Ent}_\pi(f);$$

(vi) the weak transport-entropy inequality $\tilde{T}_2^-(C)$ if for all probability measures $\nu = f\pi$, we have

$$\tilde{W}_2(\pi|f\pi)^2 \leq \frac{2}{C} \text{Ent}_\pi(f);$$

(vii) the weak transport-information inequality $\tilde{T}_2^+I(C)$ if for all probability measures $\nu = f\pi$, we have

$$\tilde{W}_2(f\pi|\pi)^2 \leq \frac{1}{C^2} \mathcal{I}_\pi(f);$$

(viii) the weak transport-information inequality $\tilde{T}_2^-I(C)$ if for all probability measures $\nu = f\pi$, we have

$$\tilde{W}_2(\pi|f\pi)^2 \leq \frac{1}{C^2} \mathcal{I}_\pi(f).$$

Here we only consider the case when the cost function is quadratic, for more general result about weak transport inequalities, we refer to [18] and [42].

The weak transport-entropy inequality $\tilde{T}_2^+(C)$ implies T_1 , so that it still implies gaussian concentration. Moreover, unlike the T_1 inequality, it has a tensorization property (see Theorem 4.11 in [18]), although a weaker form than for T_2 .

Our main results are the following theorems.

Theorem 1.10. *Let K be an irreducible Markov kernel on \mathcal{X} and π the reversible invariant probability measure associated to K . Assume that $CD(\kappa, \infty)$ holds with $\kappa > 0$. Then π satisfies the transport-information inequality $T_1 I$ with constant κ . More precisely, for all probability measure $\nu := f\pi$ on \mathcal{X} , it holds*

$$W_1(f\pi, \pi)^2 \leq \frac{2J}{\kappa^2} \mathcal{I}_\pi(f).$$

With such a result in hand, we can then follow [20] to prove a transport-entropy inequality T_1 holds, so that the Gaussian concentration property follows as well. Another application is that after a simple computation, one can obtain the following Bonnet–Myers type theorem.

Corollary 1.11. *Assume that $CD(\kappa, \infty)$ holds, then*

$$d(x, y)\kappa \leq 2(J(x) + J(y)).$$

Recall that under coarse Ricci curvature condition, [37] obtains the same type of inequality: $\kappa_c d(x, y) \leq J(x) + J(y)$.

Now if we make the stronger assumption $CDE'(\kappa_e, \infty)$, we get a stronger inequality:

Theorem 1.12. *Let (\mathcal{X}, d) be a connected graph equipped with graph distance d . Let K be a irreducible Markov kernel on \mathcal{X} and π the reversible invariant probability measure associated to K . Assume that $CDE'(\kappa_e, \infty)$ holds with $\kappa_e > 0$. Then π satisfies the transport-information inequality $\tilde{T}_2^+ I$ with constant $\kappa_e/\sqrt{2}$. More precisely, for all probability measure $\nu := f\pi$ on \mathcal{X} , it holds*

$$\tilde{W}_2(f\pi|\pi)^2 \leq \frac{2J}{\kappa_e^2} \mathcal{I}_\pi(f) \leq \frac{2}{\kappa_e^2} \mathcal{I}_\pi(f).$$

Again, following the ideas of [20], one can prove a weak-transport entropy inequality \tilde{T}_2^+ . On the other hand, sine the weak-transport cost is stronger than the L^1 -Wasserstein distance, it yields immediately $T_1 I$ holds, which implies T_1 and concentration results.

Under coarse Ricci curvature condition, the inequality $T_1 I(\kappa_c)$ holds.

Theorem 1.13. *Let \mathcal{X}, π, K define as before. If the global coarse Ricci curvature is bounded from below by $\kappa_c > 0$, then the following transport inequality holds for all density function f :*

$$W_1(f\pi, \pi)^2 \leq \frac{1}{\kappa_c^2} \mathcal{I}_\pi(f) \left(J - \frac{1}{8} \mathcal{I}_\pi(f) \right) \leq \frac{1}{\kappa_c^2} \mathcal{I}_\pi(f).$$

As a corollary, this last result implies a T_1 inequality for such Markov chain, which has been previously obtained by [13].

The paper is organized as follows: in first section we will explain why T_2 cannot hold in general, then establish connection of results in [42] and the Fisher information on graph settings. Section 2 gives a few preliminary results about Hamilton–Jacobi equations on graphs. In the third, fourth and fifth sections we will discuss functional inequalities under $CD(\kappa, \infty)$, $CDE'(\kappa_e, \infty)$ and coarse Ricci curvature κ_c respectively. In the last section we will show some applications, such as how a transport-information inequality implies a transport-entropy inequality, concentration results, a discrete analogue of the Bonnet–Meyer theorem, and a study of the example of the discrete hypercube.

2. Preliminary

In this section, we present some general results in the discrete setting, without assuming any curvature condition. Our main concern is to present the Hamilton–Jacobi equations on graphs introduced in [18,42] and their relation with weak transport costs.

As we mentioned in the introduction, in the discrete setting, when π is not a Dirac mass, the inequality $T_2(\kappa)$ cannot hold true, for any $\kappa > 0$. To our knowledge, this was first proved in [16]. We give here a different proof, as a consequence of a more general result.

Lemma 2.1. *Let (\mathcal{X}, d) be a metric space and μ a probability measure defined on \mathcal{X} . Assume that there exist $C_1, C_2 \subset \mathcal{X}$ such that*

- (i) $\inf_{x \in C_1, y \in C_2} d(x, y) > 0$,
- (ii) $\text{supp}(\mu) \subset C_1 \cup C_2$,
- (iii) $\mu(C_1) > 0, \mu(C_2) > 0$.

Then μ does not satisfies $T_2(\kappa)$ for any $\kappa > 0$.

Proof. For $h < \min\{\mu(C_1), \mu(C_2)\}$, define

$$v_h(dx) := \begin{cases} \mu(dx) \left(1 + \frac{h}{\mu(C_1)}\right), & x \in C_1, \\ \mu(dx) \left(1 - \frac{h}{\mu(C_2)}\right), & x \in C_2. \end{cases}$$

Let $d := \inf_{x \in C_1, y \in C_2} d(x, y) > 0$. Then we have $W_2(\mu, v)^2 \geq d^2 h$, and the entropy is

$$(\mu(C_1) + h) \log(1 + h/\mu(C_1)) + (\mu(C_2) - h) \log(1 - h/\mu(C_2)).$$

When h goes to 0, the entropy is $\mathcal{O}(h^2)$. The conclusion follows since W_2^2 have order $\mathcal{O}(h)$. \square

Thus, if π satisfies T_2 on graphs, it means that π is a Dirac mass.

In this section, we shall describe the links between weak transport inequalities and the Hamilton–Jacobi operator that was introduced in [18] and studied in [42].

Following [18], the weak transport cost \tilde{W}_2^2 between two probability measures μ and ν satisfies the following duality formula:

$$\tilde{W}_2(\nu|\mu)^2 = \sup_{g \in \mathcal{C}_c^b} \left\{ \int \tilde{Q}_1 g \, d\nu - \int g \, d\mu \right\}, \tag{3}$$

where the infimum-convolution operator is defined as

$$\tilde{Q}_t \varphi(x) := \inf_{p \in \mathcal{P}(\mathcal{X})} \left\{ \int g(y) p(dy) + \frac{1}{t} \left(\int d(x, y) p(dy) \right)^2 \right\}.$$

Later, the second author remarked (see [42]) that the operator \tilde{Q}_t satisfies a discrete version of the Hamilton–Jacobi equation: for all $t > 0$

$$\frac{\partial}{\partial t} \tilde{Q}_t g + \frac{1}{4} |\tilde{\nabla} \tilde{Q}_t g|^2 \leq 0, \tag{4}$$

where $|\tilde{\nabla} g|(x) := \sup_{y \in \mathcal{X}} \frac{[g(y) - g(x)]_-}{d(x, y)}$. We refer to [4,45] for information about Hamilton–Jacobi equations in the continuous setting and their link with functional inequalities.

Now let $\alpha \in \mathcal{C}_1(\mathbb{R}^+)$, according to (4), one can easily check that for all $t > 0$, it holds

$$\left| \alpha'(t) \frac{\partial}{\partial t} \tilde{Q}_{\alpha(t)} g \right| \geq \frac{1}{4} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g|^2. \tag{5}$$

The evolution with respect to time is controlled by this special “gradient”. We refer readers to [42] for properties of \tilde{Q} and $\tilde{\nabla}$. Here we shall develop some more.

Proposition 2.2 (Convexity). *Let g be a function defined on \mathcal{X} , then for all $x \in \mathcal{X}$, the function $t \mapsto \tilde{Q}_t g(x)$ is convex.*

Proof. Fix $x \in \mathcal{X}$, define $G(t) := \tilde{Q}_t g(x)$. Observe that for any $\lambda \in [0, 1]$, and $p_1, p_2 \in \mathcal{P}(\mathcal{X})$, setting $p := \lambda p_1 + (1 - \lambda) p_2 \in \mathcal{P}(\mathcal{X})$ and applying the Cauchy–Schwarz inequality, it holds for all $t, s > 0$:

$$\begin{aligned} \left(\int d(x, z) p(dz) \right)^2 &= \left(\lambda \int d(x, z) p_1(dz) + (1 - \lambda) \int d(x, z) p_2(dz) \right)^2 \\ &\leq (\lambda t + (1 - \lambda)s) \left(\frac{\lambda (\int d(x, z) p_1(dz))^2}{t} \right. \\ &\quad \left. + \frac{(1 - \lambda) (\int d(x, z) p_2(dz))^2}{s} \right). \end{aligned} \tag{6}$$

As a consequence, we get

$$\begin{aligned} & \lambda \left(\int g(z) p_1(dz) + \frac{1}{t} \left(\int d(x, z) p(dz) \right)^2 \right) \\ & + (1 - \lambda) \left(\int g(z) p_2(dz) + \frac{1}{s} \left(\int d(x, z) p_2(dz) \right)^2 \right) \\ & \geq \int g(z) p(dz) + \frac{1}{(\lambda t + (1 - \lambda)s)} \left(\int d(x, z) p(dz) \right)^2 \geq G(\lambda t + (1 - \lambda)s). \end{aligned} \tag{7}$$

Taking the infimum over all $p_1, p_2 \in \mathcal{P}(\mathcal{X})$ on left hand side of the inequality, the conclusion follows. \square

The following lemma is a technical result connecting the gradient $\tilde{\nabla}$ and Γ -operator.

Lemma 2.3. *Let π be the reversible probability measure for the Markov kernel k . For any bounded function f and g on \mathcal{X} , the following inequalities hold:*

- (i) $\int \Gamma(f, g) d\pi \leq \sqrt{2J} \int |\tilde{\nabla} g| |\tilde{\nabla} f| d\pi,$
- (ii) $|\int \Gamma(f, g) d\pi| \leq \sqrt{2J} \int |\tilde{\nabla} g| \sqrt{\Gamma(f)} d\pi.$

Moreover, if we suppose that f is non negative, then we have

- (iii) $\int \Gamma(f, g) d\pi \leq 2\sqrt{2J} \int |\tilde{\nabla} g| \sqrt{f\Gamma(\sqrt{f})} d\pi,$
- (iv) $\int \Gamma(\sqrt{f}) d\pi \leq \frac{1}{4} \int |\tilde{\nabla} \log f|^2 f d\pi.$

Proof. The proofs of these four inequalities all follow similar arguments. Denote the positive part and negative part of a function u as u_+ and u_- , respectively.

(i): Using the relation $(uv)_+ \leq u_+v_+ + u_-v_-$, we have

$$\begin{aligned} \int \Gamma(f, g)_+ d\pi &= \frac{1}{2} \sum_x \left[\sum_{y \sim x} (f(y) - f(x))(g(y) - g(x)) \right]_+ K(x, y)\pi(x) \\ &\leq \frac{1}{2} \sum_x \sum_{y \sim x} (g(y) - g(x))_+ (f(y) - f(x))_+ K(x, y)\pi(x) \\ &\quad + \frac{1}{2} \sum_x \sum_{y \sim x} (g(y) - g(x))_- (f(y) - f(x))_- K(x, y)\pi(x). \end{aligned}$$

Now by reversibility of the measure π , it holds

$$\begin{aligned} & \sum_x \sum_x (g(y) - g(x))_+ (f(y) - f(x))_+ K(x, y)\pi(x) \\ &= \sum_x \sum_{y \sim x} (g(y) - g(x))_- (f(y) - f(x))_- K(x, y)\pi(x) \\ &\leq \sum_x |\tilde{\nabla} g|(x) \sum_{y \sim x} (f(x) - f(y))_- K(x, y)\pi(x). \end{aligned}$$

Where the latter inequality follows from $|\tilde{\nabla}g|(x) \geq (g(y) - g(x))_-$ for all $y \sim x$. Therefore, we get

$$\int \Gamma(f, g)_+ d\pi \leq \sum_x |\tilde{\nabla}g| \sum_{y \sim x} (f(y) - f(x))_- K(x, y) \pi(x). \tag{8}$$

In (8), using $\Gamma(f, g) \leq \Gamma(f, g)_+$ and $\sum_{y \sim x} (f(y) - f(x))_- K(x, y) \leq |\tilde{\nabla}f|(x)J(x)$, we get (i).

(ii): By the Cauchy–Schwarz inequality, it holds

$$\left(\sum_{y \sim x} (f(y) - f(x))_- K(x, y) \right)^2 \leq J(x) \sum_{y \sim x} (f(y) - f(x))_-^2 K(x, y) \leq 2J\Gamma(f). \tag{9}$$

Combining (8) and (9) leads to

$$\int \Gamma(f, g)_+ d\pi \leq \int |\tilde{\nabla}g| \sqrt{2J\Gamma(f)} d\pi. \tag{10}$$

Following a similar argument, we have

$$\int \Gamma(f, g)_- d\pi \leq \int |\tilde{\nabla}g| \sqrt{2J\Gamma(f)} d\pi, \tag{11}$$

and (ii) follows by (10), (11) and the inequality

$$\left| \int \Gamma(f, g) d\pi \right| \leq \max \left\{ \int \Gamma(f, g)_+ d\pi, \int \Gamma(f, g)_- d\pi \right\}.$$

(iii): Since f is nonnegative, \sqrt{f} is well defined. Then it holds

$$\begin{aligned} \int \Gamma(f, g) d\pi &= \frac{1}{2} \sum_x \sum_{y \sim x} (f(y) - f(x))(g(y) - g(x))K(x, y)\pi(x) \\ &= \frac{1}{2} \sum_x \sum_{y \sim x} (g(y) - g(x))(\sqrt{f}(y) - \sqrt{f}(x))(\sqrt{f}(y) + \sqrt{f}(x))K(x, y)\pi(x). \end{aligned}$$

Now arguing as in (i) and (ii), by reversibility of π , we get

$$\int \Gamma(f, g) d\pi = \sum_x \sum_{y \sim x} (g(y) - g(x))_-(\sqrt{f}(y) - \sqrt{f}(x))_-(\sqrt{f}(y) + \sqrt{f}(x))K(x, y)\pi(x).$$

Notice that $(\sqrt{f}(y) - \sqrt{f}(x))_-(\sqrt{f}(y) + \sqrt{f}(x)) \leq (\sqrt{f}(y) - \sqrt{f}(x))_- 2\sqrt{f}(x)$, we have

$$\begin{aligned} \int \Gamma(f, g) d\pi &\leq 2 \sum_x \sum_{y \sim x} (g(y) - g(x))_-(\sqrt{f}(y) - \sqrt{f}(x))_- \sqrt{f}(x)K(x, y)\pi(x) \\ &\leq 2\sqrt{2J} \int |\tilde{\nabla}g| \sqrt{f\Gamma(\sqrt{f})} d\pi, \end{aligned}$$

where the last step we have used (9) with $u := \sqrt{f}$.

(iv): If f is the null function, there is nothing to say. Otherwise, if there exist $x, y \in \mathcal{X}$ such that $f(x) = 0, f(y) > 0$, it is easy to see that $|\nabla \log f(y)|^2 f(y) \pi(y) = \infty$. So we only need to prove the case $f(x) > 0$ for all $x \in \mathcal{X}$.

Since f is a positive function, one can rewrite $f = e^g$, it is enough to prove that

$$\int \Gamma(e^{g/2}) d\pi \leq \frac{J}{4} \int |\tilde{\nabla} g|^2 e^g d\pi$$

holds for all function g . In fact, by convexity of function $x \mapsto e^x$, we have for all $a > b, (a - b)e^a \geq e^a - e^b$. Thus,

$$\begin{aligned} J \int |\tilde{\nabla} g|^2 e^g &\geq \sum_{x \sim y; g(y) \leq g(x)} (g(x) - g(y))^2 e^{g(x)} K(x, y) \pi(x) \\ &\geq 4 \sum_{x \sim y; g(y) \leq g(x)} \left(e^{\frac{g(x)}{2}} - e^{\frac{g(y)}{2}} \right)^2 K(x, y) \pi(x) \\ &= 4 \int \Gamma(e^{g/2}) d\pi. \end{aligned} \quad \square$$

2.1. Transport-information inequalities implies transport-entropy inequalities

We prove here the discrete version of Theorem 2.1 in [20], which states that we can deduce transport-entropy inequalities from transport-information inequalities. The proof is essentially unchanged, we give it to justify the validity of the theorem in the discrete setting.

Theorem 2.4. *Assume that the transport-information inequality*

$$W_1(f\pi, \pi)^2 \leq \frac{1}{C^2} \mathcal{I}_\pi(f)$$

holds. Then we have the transport-entropy inequality

$$W_1(f\pi, \pi)^2 \leq \frac{2}{C} \text{Ent}_\pi(f).$$

Proof. The transport-entropy inequality $T_1(C)$ is equivalent to the estimate

$$\int e^{\lambda f} d\pi \leq \exp\left(\frac{\lambda^2}{2C}\right)$$

for all 1-Lipschitz function f with $\int f d\pi = 0$ and all $\lambda \geq 0$. Let f be such a function. Let $Z(\lambda) := \int e^{\lambda f} d\pi$ and $\mu_\lambda := e^{\lambda f} \pi / Z(\lambda)$. We have

$$\frac{d}{d\lambda} \log Z(\lambda) = \frac{1}{Z(\lambda)} \int f e^{\lambda f} d\pi \leq W_{1,d}(\mu_\lambda, \pi) \leq \sqrt{\frac{4}{C^2} \int \frac{\Gamma(e^{\lambda f/2})}{Z(\lambda)} d\pi}.$$

Using the inequality $\int \Gamma(f) d\pi \leq \int f^2 \Gamma(\log f) d\pi$, we deduce

$$\frac{d}{d\lambda} \log Z(\lambda) \leq \frac{\lambda}{C}$$

which integrates into $\log Z(\lambda) \leq \lambda^2/(2C)$, and this is the bound we were looking for. □

We shall now show that the weak transport-information inequality $\tilde{T}_2^+ I$ implies the weak-transport-entropy inequality \tilde{T}_2^+ . The proof is an adaptation of the one for the T_2 and $T_2 I$ inequalities in the continuous setting from [20].

Theorem 2.5. *Assume that π satisfies the modified weak-transport information inequality $\tilde{T}_2^+ I(C)$, then π satisfies the weak-transport inequality $\tilde{T}_2^+(C)$.*

Proof. According to [18], the transport-entropy inequality \tilde{T}_2 is equivalent to

$$\int \exp\left(\frac{2}{C} \tilde{Q}_1 f\right) d\pi \leq \exp\left(\frac{2}{C} \int f d\pi\right) \quad \forall f : \mathcal{X} \rightarrow \mathbb{R} \text{ bounded.}$$

Usually, the class of functions f we must use is the class of bounded continuous functions, but here, since we work on a discrete space endowed with a graph distance, we only have to work with bounded functions.

We write $F(t) := \log \int \exp(k(t) \tilde{Q}_t f) d\pi - k(t) \int f d\pi$ with $k(t) := Ct$. Let μ_t be the probability measure with density with respect to π proportional to $\exp(k(t) \tilde{Q}_t f)$.

According to part (iv) of Lemma 2.3, we have

$$\int \Gamma(e^{\frac{1}{2}(k(t) \tilde{Q}_t f)}) d\pi \leq \int |\tilde{\nabla} k(t) \tilde{Q}_t f|^2 d\mu_t.$$

Hence, for $t > 0$, we have

$$\begin{aligned} F'(t) &\leq \frac{1}{\int \exp(k(t) \tilde{Q}_t f) d\pi} \left(\int k'(t) \tilde{Q}_t f e^{k(t) \tilde{Q}_t f} d\pi - \int k(t) |\tilde{\nabla} \tilde{Q}_t f|^2 e^{k(t) \tilde{Q}_t f} d\pi \right) \\ &\quad - k'(t) \int f d\pi \\ &\leq \frac{k'(t)}{t} \left(\int \tilde{Q}_1(tf) d\mu_t - \int tf d\pi \right) - k(t) \int |\tilde{\nabla} \tilde{Q}_t f|^2 d\mu_t \\ &\leq \frac{k'(t)}{t} \tilde{W}_2(\mu_t, \pi)^2 - k(t) \int |\tilde{\nabla} \tilde{Q}_t f|^2 d\mu_t \\ &\leq \frac{k'(t)}{tC^2} \int \Gamma(e^{\frac{1}{2}(k(t) \tilde{Q}_t f)}) d\pi - k(t) \int |\tilde{\nabla} \tilde{Q}_t f|^2 d\mu_t \\ &\leq \left(\frac{k'(t)}{tC^2} - \frac{1}{k(t)} \right) \int |\tilde{\nabla} k(t) \tilde{Q}_t f|^2 d\mu_t \\ &= 0. \end{aligned}$$

□

3. Transport inequalities for Markov chains satisfying $CD(\kappa, \infty)$

In this section, we assume that the Markov chain satisfies the curvature condition $CD(\kappa, \infty)$ for some $\kappa > 0$. One of the main tools we shall use is the following sub-commutation relation between Γ and the semigroup P_t , which was obtained in [26].

Lemma 3.1. *Assume that $CD(\kappa, \infty)$ holds. Then for any $f : \mathcal{X} \rightarrow \mathbb{R}$, we have*

$$\Gamma(P_t f) \leq e^{-2\kappa t} P_t \Gamma(f).$$

Another property we shall make repeated use of is that if a function is 1-Lipschitz with respect to d , then automatically $\Gamma(f) \leq J$. This is an immediate consequence of the definitions of Γ and of the graph distance d adapted to the Markov kernel K .

Proof of Theorem 1.10. The proof relies on the Kantorovitch–Rubinstein duality formula

$$W_1(\pi, \nu) = \sup_{g \text{ 1-lip}} \int g d\pi - \int g d\nu.$$

Let g be a 1-Lipschitz function. It implies that $\Gamma(g) \leq J$.

First, using the Cauchy–Schwarz inequality, it holds

$$\begin{aligned} & - \int \Gamma(P_t g, f) d\pi \\ &= -\frac{1}{2} \sum_{x,y} (P_t g(y) - P_t g(x))(f(y) - f(x)) K(x, y) \pi(x) \\ &= \frac{1}{2} \sum_{x,y} |(P_t g(y) - P_t g(x))(\sqrt{f}(y) - \sqrt{f}(x))(\sqrt{f}(y) + \sqrt{f}(x))| K(x, y) \pi(x) \\ &\leq \sum_x \pi(x) \Gamma(\sqrt{f})(x)^{\frac{1}{2}} \\ &\quad \times \left(\sum_y (P_t g(y) - P_t g(x))^2 (\sqrt{f}(y) + \sqrt{f}(x))^2 K(x, y) \right)^{\frac{1}{2}}. \end{aligned}$$

Now applying the Cauchy–Schwarz inequality again, the latter quantity is less than

$$\begin{aligned} & \left(\int \Gamma(\sqrt{f}) d\pi \right)^{\frac{1}{2}} \\ & \quad \times \left(\sum_{x,y} (P_t g(y) - P_t g(x))^2 (\sqrt{f}(y) + \sqrt{f}(x))^2 K(x, y) \pi(x) \right)^{\frac{1}{2}}. \end{aligned}$$

Therefore, we have

$$\begin{aligned}
 & - \int \Gamma(P_t g, f) d\pi \\
 & \leq \left(\int \Gamma(\sqrt{f}) d\pi \right)^{\frac{1}{2}} \left(\sum_{x,y} (P_t g(y) - P_t g(x))^2 (\sqrt{f}(y) + \sqrt{f}(x))^2 K(x, y) \pi(x) \right)^{\frac{1}{2}} \\
 & \leq \sqrt{2} \sqrt{\mathcal{I}_\pi(f)} \sqrt{\int \Gamma(P_t g) f d\pi},
 \end{aligned}$$

where the last step we have used the reversibility of the measure π and the fact that

$$(\sqrt{f}(y) + \sqrt{f}(x))^2 \leq 2(f(x) + f(y)) \tag{12}$$

for any nonnegative function f .

Therefore, according to Lemma 3.1, we have

$$\begin{aligned}
 \int g d\pi - \int g f d\pi &= \int_0^{+\infty} \frac{d}{dt} \int (P_t g) f d\pi dt \\
 &= - \int_0^{+\infty} \int \Gamma(P_t g, f) d\pi dt \\
 &\leq \int_0^{+\infty} \sqrt{\mathcal{I}_\pi(f)} \sqrt{\int \Gamma(P_t g) f d\pi} dt \\
 &\leq \sqrt{2} \sqrt{\mathcal{I}_\pi(f)} \int_0^{+\infty} e^{-\kappa t} \sqrt{\int P_t(\Gamma(g)) f d\pi} dt \\
 &\leq \frac{\sqrt{2J}}{\kappa} \sqrt{\mathcal{I}_\pi(f)}.
 \end{aligned}$$

The result immediately follows by taking the supremum over all 1-Lipschitz functions g . □

Using, similar arguments, we can also prove the following Cheeger-type inequality.

Proposition 3.2. *Assume that $CD(\kappa, \infty)$ holds. Then for any probability density f with respect to π , we have*

$$W_1(f\pi, \pi) \leq \frac{\sqrt{J}}{\kappa} \int \sqrt{\Gamma(f)} d\pi.$$

We call this a Cheeger-type inequality, by analogy with the classical Cheeger inequality

$$\|f\pi - \pi\|_{TV} \leq C \int |\nabla f| d\pi.$$

Here $\int \sqrt{\Gamma(f)} d\pi$ is an L^1 estimate on the gradient of f , while $\|f\pi - \pi\|_{TV}$ is also a distance of L^1 nature.

Proof. Once more, by Kantorovitch duality for W_1 , and since 1-Lipschitz functions g satisfy $\Gamma(g) \leq J$, we have

$$\begin{aligned} W_1(f\pi, \pi) &\leq \sup_{g; \Gamma(g) \leq J} \int gf d\pi - \int g d\pi \\ &= \sup_{g; \Gamma(g) \leq J} - \int_0^{+\infty} \int \Gamma(P_t g, f) d\pi dt \\ &\leq \sup_{g; \Gamma(g) \leq J} \int_0^{+\infty} \int \sqrt{\Gamma(P_t g)} \sqrt{\Gamma(f)} d\pi dt \\ &\leq \sup_{g; \Gamma(g) \leq J} \int_0^{+\infty} e^{-\kappa t} \int \sqrt{P_t \Gamma(g)} \sqrt{\Gamma(f)} d\pi dt \\ &\leq \frac{\sqrt{J}}{\kappa} \int \sqrt{\Gamma(f)} d\pi. \end{aligned} \quad \square$$

4. Transport inequalities for Markov chain satisfying $CDE'(\kappa_e, \infty)$

In this section, we assume that the exponential curvature condition $CDE'(\kappa_e, \infty)$ holds. We will prove Theorem 1.12. But first, we shall study some properties of the $CDE'(\kappa_e, \infty)$ condition.

4.1. Properties of $CDE'(\kappa_e, \infty)$

Lemma 4.1. *Assume that $CDE'(\kappa, \infty)$ holds. Then for any nonnegative function $f : \mathcal{X} \rightarrow \mathbb{R}$ and any $t \geq 0$, we have*

- (i) $\Gamma(\sqrt{P_t f}) \leq e^{-2\kappa_e t} P_t \Gamma(\sqrt{f})$.
- (ii) $\frac{\Gamma(P_t f)}{P_t f} \leq e^{-2\kappa_e t} P_t \left(\frac{\Gamma(f)}{f}\right)$.

Remark 4.1. (i) looks like the commutation formula of Γ and $\sqrt{\cdot}$ under $CD(\kappa, \infty)$ in continuous settings. But it is not the same thing, the positivity is very important. Recall that in classical Bakry–Émery theory, the commutation formula is the following: for all f (smooth enough), $\sqrt{\Gamma(P_t f)} \leq e^{-\kappa t} P_t(\sqrt{\Gamma(f)})$. We have not been able to recover this formula under $CDE'(\kappa_e, \infty)$ in graphs settings.

Proof. The proof follows a standard interpolation argument. We begin with (i). Let $g := P_{t-s} f$ and define $\varphi(s) := e^{-2\kappa_e s} P_s(\Gamma(\sqrt{g}))$. To obtain the result, it is enough to show that $\varphi' \geq 0$. In

fact,

$$\begin{aligned} \varphi'(s) &= e^{-2\kappa_e s} P_s \left[L(\Gamma(\sqrt{g})) - \Gamma\left(\sqrt{g}, \frac{Lg}{\sqrt{g}}\right) - 2\kappa_e \Gamma(\sqrt{g}) \right] \\ &\geq 0, \end{aligned}$$

where we have used the assumption on the curvature, which is equivalent to

$$\frac{1}{2}L\Gamma(f) - \Gamma\left(f, \frac{L(f^2)}{2f}\right) \geq \kappa_e \Gamma(f)$$

(see (3.11) in [5]).

Similarly, let $\psi(s) := e^{-2\kappa_e s} P_s(\frac{\Gamma(g)}{g})$. Again, it is enough to show that $\psi' \geq 0$. We have

$$\psi'(s) = e^{-\kappa_e s} P_s \left(L\left(\frac{\Gamma(g)}{g}\right) + \frac{1}{g^2}(-2g\Gamma(g, Lg) + \Gamma(g)Lg) - 2\kappa_e \frac{\Gamma(g)}{g} \right). \tag{13}$$

Since g is positive, we only need to show that

$$g \left(L\left(\frac{\Gamma(g)}{g}\right) + \frac{1}{g^2}(-2g\Gamma(g, Lg) + \Gamma(g)Lg) - 2\kappa_e \frac{\Gamma(g)}{g} \right) \geq 0. \tag{14}$$

Notice that (14) is equivalent to

$$gL\left(\frac{\Gamma(g)}{g}\right) - 2\Gamma(g, Lg) + \frac{1}{g}\Gamma(g)Lg \geq 2\kappa_e \Gamma(g),$$

and we conclude by writing

$$2\kappa_e \Gamma(g) \leq 2\tilde{\Gamma}_2(g) = 2\left(\Gamma_2(g) - \Gamma\left(g, \frac{\Gamma(g)}{g}\right)\right) = gL\left(\frac{\Gamma(g)}{g}\right) - 2\Gamma(g, Lg) + \frac{1}{g}\Gamma(g)Lg. \quad \square$$

4.2. Weak transport-information inequalities under CDE'(κ_e, ∞)

Using Lemma 4.1, we can prove some weak transport-information inequalities under CDE'(κ_e, ∞). First, we will prove Theorem 1.12.

Proof of Theorem 1.12. Let $\alpha(t) = e^{-\kappa_e t}$, for any probability density f with respect to π and any $t > 0$, applying (5), it holds:

$$\begin{aligned} &\int_0^\infty \frac{d}{dt} \int \tilde{Q}_{\alpha(t)g} P_t f \, d\pi \, dt \\ &\leq \int_0^\infty \int -\frac{\kappa_e}{4} e^{-\kappa_e t} |\tilde{\nabla} \tilde{Q}_{\alpha(t)g}|^2 P_t f + \Gamma(\tilde{Q}_{\alpha(t)g}, P_t f) \, d\pi \, dt. \end{aligned} \tag{15}$$

According to part (iii) of Lemma 2.3, we have

$$\begin{aligned} & \int \tilde{Q} g f \, d\pi - \int g \, d\pi \\ &= \int_0^\infty \frac{d}{dt} \int \tilde{Q}_{\alpha(t)} g P_t f \, d\pi \, dt \\ &\leq \int_0^\infty \int -\frac{\kappa_e}{4} e^{-\kappa_e t} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g|^2 P_t f + 2\sqrt{2J} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g| \sqrt{P_t f \Gamma(\sqrt{P_t f})} \, d\pi \, dt \\ &\leq \int_0^\infty \frac{8J e^{\kappa_e t}}{\kappa_e} \int \Gamma(\sqrt{P_t f}) \, d\pi \, dt. \end{aligned}$$

Now we apply Lemma 4.1, and we get

$$\begin{aligned} \int \tilde{Q} g f \, d\pi - \int g \, d\pi &\leq \int_0^\infty \frac{8e^{-\kappa_e t}}{\kappa_e} \int P_t(\Gamma(\sqrt{f})) \, d\pi \, dt \\ &= \frac{8J}{\kappa_e^2} \int \Gamma(\sqrt{f}) \, d\pi = \frac{2J}{\kappa_e^2} \mathcal{I}_\pi(f). \end{aligned}$$

The conclusion then follows from the duality formula (3) by taking $\mu = \pi$ and $\nu = f\pi$. \square

It is easy to see that $\mathcal{I}_\pi(f) \leq \bar{\mathcal{I}}_\pi(f) := \int \frac{\Gamma(f)}{f} \, d\pi$, thus we have the following corollary.

Corollary 4.2. *Assume that the exponential curvature condition $\text{CDE}'(\kappa_e, \infty)$ holds, then π satisfies the following weak-transport information inequalities:*

$$\tilde{W}_2(f\pi | \pi)^2 \leq \frac{2J}{\kappa_e^2} \bar{\mathcal{I}}_\pi(f).$$

Unfortunately, we have not been able to establish the relation of $\tilde{W}_2(\pi | f\pi)^2$ and $\int \Gamma(\sqrt{f}) \, d\pi$. However, as in Corollary 4.2, we get a weaker inequality as follows:

Theorem 4.3. *Assume that the exponential curvature condition $\text{CDE}'(\kappa_e, \infty)$ holds, then π satisfies the following weak-transport information inequalities:*

$$\tilde{W}_2(\pi | f\pi)^2 \leq \frac{2J}{\kappa_e^2} \bar{\mathcal{I}}_\pi(f).$$

Proof. We prove this theorem in a similar way as the previous one.

Let $\alpha(t) := 1 - e^{-\kappa t}$. Arguing as in the latter theorem, we get

$$\begin{aligned} \int \tilde{Q} g \, d\pi - \int g f \, d\pi &= \int_0^\infty \frac{d}{dt} \int \tilde{Q}_{\alpha(t)} g P_t f \, d\pi \, dt \\ &\leq \int_0^\infty \int -\frac{\kappa}{4} e^{-\kappa t} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g|^2 P_t f - \Gamma(\tilde{Q}_{\alpha(t)} g, P_t f) \, d\pi \, dt. \end{aligned}$$

Now applying part (ii) of Lemma 2.3, it follows that

$$\begin{aligned} & \int \tilde{Q}g \, d\pi - \int gf \, d\pi \\ & \leq \int_0^\infty \int -\frac{\kappa_e}{4} e^{-\kappa_e t} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g|^2 P_t f + \int |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g| \sqrt{2J\Gamma(P_t f)} \, d\pi \, dt \\ & \leq \int_0^\infty \frac{2J e^{\kappa_e t}}{\kappa_e} \int \frac{\Gamma(P_t f)}{P_t f} \, d\pi \, dt \leq \int_0^\infty \frac{2J e^{-\kappa_e t}}{\kappa_e} \int P_t \left(\frac{\Gamma(f)}{f} \right) \, d\pi \, dt \\ & = \frac{2J}{\kappa_e^2} \int \frac{\Gamma(f)}{f} \, d\pi. \end{aligned} \quad \square$$

Remark 4.2. (i) One can get Corollary 4.2 by a similar argument: let $\alpha(t) := e^{-\kappa t}$

$$\begin{aligned} \int \tilde{Q}gf \, d\pi - \int g \, d\pi &= \int_0^\infty \frac{d}{dt} \int \tilde{Q}_{\alpha(t)} g P_t f \, d\pi \, dt \\ &\leq \int_0^\infty \int -\frac{\kappa}{2} e^{-\kappa t} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g|^2 P_t f + \sqrt{2J} |\tilde{\nabla} \tilde{Q}_{\alpha(t)} g| \sqrt{|\Gamma(P_t f)|} \, d\pi \, dt \\ &\leq \int_0^\infty \frac{J e^{\kappa t}}{\kappa_e} \int \frac{\Gamma(P_t f)}{P_t f} \, d\pi \, dt \leq \frac{J}{\kappa_e^2} \int \frac{\Gamma(f)}{f} \, d\pi \, dt. \end{aligned}$$

(ii) Using the notations of [18], define $\tilde{W}_2(f\pi, \pi)^2 = \frac{1}{2}(\tilde{W}_2^2(f\pi|\pi) + \tilde{W}_2^2(\pi|f\pi))$, denote $\mathcal{P}_2(\mathcal{X})$ as the set of the probability measure on \mathcal{X} which has a finite second moment. Then $(\mathcal{P}_2(\mathcal{X}), \tilde{W}_2(\cdot, \cdot))$ is a metric space, and if \mathcal{X} satisfies the exponential curvature condition, we have an upper bound for $\tilde{W}_2(\cdot, \pi)$ in terms of modified Fisher information. Of course, when we work on a finite space, any probability measure has finite second moment.

5. Transport-information inequality for Markov chains with positive coarse Ricci curvature

In this section, we assume the Markov chain has coarse Ricci curvature bounded from below by κ_c , with respect to the graph distance d .

As a consequence of the bound on the curvature, note that for any 1-Lipschitz function g , $P_t g$ is $e^{-\kappa_e t}$ -Lipschitz.

The problem of proving a transport-entropy inequality for Markov chains with positive coarse Ricci curvature was raised in Problem J in [38]. It was proved by [13]. The transport-information inequality is a slight improvement of this result. Note that T_1 cannot hold in the full generality of the setting of [37], since it implies Gaussian concentration, which does not hold for some examples with positive curvature, more details and examples can be found in [37].

The proof of this result will make use of the following lemma.

Lemma 5.1. *If the coarse Ricci curvature is bounded from below by $\kappa_c > 0$, then*

$$W_1(f\pi, \pi) \leq \frac{1}{\kappa_c} \sum_{x \neq y} |f(x) - f(y)| K(x, y) \pi(x).$$

Proof. By Kantorovitch duality for W_1 , we have

$$\begin{aligned} W_1(f\pi, \pi) &= \sup_{g \text{ 1-lip}} \int g f d\pi - \int g d\pi = \sup_{g \text{ 1-lip}} - \int_0^{+\infty} \frac{d}{dt} \int P_t g f d\pi dt \\ &= - \int_0^{+\infty} \sum_{x, y} (P_t f(y) - P_t g(x))(f(y) - f(x)) K(x, y) \pi(x) dt \\ &\leq \int_0^{+\infty} \|P_t g\|_{\text{lip}} \sum_{x, y} |f(y) - f(x)| K(x, y) \pi(x) dt \\ &\leq \frac{1}{\kappa_c} \sum_{x \neq y} |f(x) - f(y)| K(x, y) \pi(x). \end{aligned}$$

□

We can now prove Theorem 1.13.

Proof of Theorem 1.13. Observe that

$$\begin{aligned} &\sum_{x \neq y} (\sqrt{f}(x) + \sqrt{f}(y))^2 K(x, y) \pi(x) \\ &= \sum_{x \neq y} (2f(x) + 2f(y) - (\sqrt{f}(x) - \sqrt{f}(y))^2) K(x, y) \pi(x) \\ &\leq \sum_{x \neq y} (2f(x) + 2f(y)) K(x, y) \pi(x) - \sum_{x \neq y} (\sqrt{f}(x) - \sqrt{f}(y))^2 K(x, y) \pi(x) \\ &\leq 4J - \frac{1}{2} \mathcal{I}_\pi(f). \end{aligned}$$

Now using Lemma 5.1, we have

$$\begin{aligned} W_1(f\pi, \pi) &\leq \frac{1}{\kappa_c} \sum_{x \neq y} |f(x) - f(y)| K(x, y) \pi(x) \\ &= \frac{1}{\kappa_c} \sum_{x \neq y} |\sqrt{f}(x) - \sqrt{f}(y)| (\sqrt{f}(x) + \sqrt{f}(y)) K(x, y) \pi(x) \\ &\leq \frac{1}{\kappa_c} \sqrt{\mathcal{I}_\pi(f)} \sqrt{\frac{1}{4} \sum_{x \neq y} (\sqrt{f}(x) + \sqrt{f}(y))^2 K(x, y) \pi(x)} \\ &\leq \frac{1}{\kappa_c} \sqrt{\mathcal{I}_\pi(f)} \sqrt{J - \frac{1}{8} \mathcal{I}_\pi(f)}. \end{aligned}$$

□

6. Applications

6.1. Transport-information inequalities imply diameter bounds

We now show that transport-information inequalities imply diameter bounds, in the spirit of the L^1 Bonnet–Myers theorem of [37]. The main idea is the observation following proposition.

Proposition 6.1. *If $\nu = f\pi$ is a dirac measure, say δ_z , then it holds*

$$\mathcal{I}_\pi(f) = 4J(z).$$

Proof. Since f is the density function correspond to the Dirac mass $\nu = \delta_z$, $f := \frac{1_z}{\pi(z)}$.

$$\begin{aligned} \int \Gamma(\sqrt{f}) d\pi &= \frac{1}{2} \left(\sum_{y \sim z} (\sqrt{f(z)})^2 k(z, y)\pi(z) + \sum_{y \sim z} (\sqrt{f(z)})^2 k(y, z)\pi(y) \right) \\ &= \sum_{y \sim z} (\sqrt{f(z)})^2 k(z, y)\pi(z) = J(z). \end{aligned}$$

□

With this proposition, one deduce immediately the following theorem.

Theorem 6.2 (Diameter estimate). *Assume that the transport-information inequality*

$$W_{1,d}(f\pi, \pi)^2 \leq \frac{1}{C^2} \mathcal{I}_\pi(f)$$

holds, for some distance d . Then

$$\sup_{x, y \in \mathcal{X}} d(x, y) \leq \frac{2}{C} (\sqrt{J(x)} + \sqrt{J(y)}).$$

Remark 6.1. Here d can be any distance defined on \mathcal{X} .

Proof. Let f_x and f_y be the density functions of δ_x and δ_y with respect to π . According to Proposition 6.1, it holds for all $x, y \in \mathcal{X}$,

$$d(x, y) = W_1(\delta_x, \delta_y) \leq W_1(f_x\pi, \pi) + W_1(f_y\pi, \pi) \leq \frac{1}{C} (\sqrt{4J(x)} + \sqrt{4J(y)}). \quad \square$$

Since $CD(\kappa, \infty)$ implies the L^1 transport information inequality, combining with the latter theorem, one can obtain a diameter estimate which is weaker than Corollary 1.11. In order to obtain Corollary 1.11, we need to revisit the proof of theorem 1.10 and prove the following lemma.

Proposition 6.3. *Assume that $CD(\kappa, \infty)$ holds, then $W_1(\pi, \delta_z) \leq \frac{2J(z)}{\kappa}$.*

The proof is essentially the same as the proof of theorem 1.10, the fact that supposing $\nu = \delta_z$ leads to some better constants.

Proof. Denote f the density function of δ_z with respect to π . Observe that f satisfies

$$(\sqrt{f}(y) + \sqrt{f}(x))^2 \leq (f(x) + f(y)). \tag{16}$$

Then following the lines of the proof of Theorem 1.10, we have for all 1-Lipschitz function g ,

$$\begin{aligned} & - \int \Gamma(P_t g, f) d\pi \\ & \leq \left(\int \Gamma(\sqrt{f}) d\pi \right)^{\frac{1}{2}} \left(\sum_{x,y} (P_t g(y) - P_t g(x))^2 (\sqrt{f}(y) + \sqrt{f}(x))^2 K(x, y) \pi(x) \right)^{\frac{1}{2}} \\ & \leq \sqrt{\mathcal{I}_\pi(f)} \sqrt{\int \Gamma(P_t g) f d\pi} \\ & = \sqrt{\mathcal{I}_\pi(f)} \sqrt{\Gamma(P_t g)(z)}. \end{aligned}$$

Applying Proposition 6.1 and inequality $\Gamma(P_t g) \leq e^{-2\kappa t} \Gamma(g)$, one deduce that

$$- \int \Gamma(P_t g, f) d\pi \leq 2\sqrt{J(z)\Gamma(P_t g)(z)} \leq 2e^{-\kappa t} J(z).$$

Therefore, according to Lemma 3.1, we have

$$\begin{aligned} \int g d\pi - \int g f d\pi &= \int_0^{+\infty} \frac{d}{dt} \int (P_t g) f d\pi dt \\ &= - \int_0^{+\infty} \int \Gamma(P_t g, f) d\pi dt \\ &\leq \frac{2J(z)}{\kappa}. \end{aligned}$$

The proof is completed. □

Now we are ready to prove Corollary 1.11

Proof of Corollary 1.11. According to the latter proposition, we have

$$d(x, y) = W_1(\delta_x, \delta_y) \leq W_1(\delta_x, \pi) + W_1(\pi, \delta_y) \leq \frac{2}{\kappa} (J(x) + J(y)). \tag{□}$$

If we look at the example of the discrete hypercube of dimension N (see the next subsection), with our notations it satisfies $CD(1/N, \infty)$. The above theorem gives the correct bound on the diameter for the graph distance of N .

6.2. An example: The discrete hypercube

As an example of Markov chain satisfying $\text{CDE}'(\kappa, \infty)$, we study the example of the symmetric random walk on the discrete hypercube. It is a Markov chain on $\{0, 1\}^N$, which at rate 1 selects a coordinate uniformly at random, and flips it with probability $1/2$. The transition rates are $K(x, y) = 1/(2N)$ for x, y such that $d(x, y) = 1$, and else it is 0.

Theorem 6.4. *The symmetric random walk on the discrete hypercube satisfies $\text{CDE}'(1/N, \infty)$.*

Proof. We start with the case $N = 1$. Since then we only have to consider a Markov chain on a two-points space, we can easily do explicit computations. Fix $f : \{0, 1\} \rightarrow \mathbb{R}$. We have

$$\Gamma(f)(0) = \Gamma(f)(1) = \frac{1}{4}(f(0) - f(1))^2$$

and hence $\Gamma(f, \frac{\Gamma(f)}{f}) = 0$ and

$$\tilde{G}_2(f) = \Gamma_2(f) = -\Gamma(f, Lf) = \Gamma(f).$$

Therefore when $N = 1$, the Markov chain satisfies $\text{CDE}'(1, \infty)$.

The general case follows, using a tensorization argument. In the unnormalized case, using Proposition 3.3 of [6], the graph satisfies $\text{CDE}'(1, \infty)$ independently of N . Since we consider the case of a Markov chain and enforce (1), we rescale the generator by a factor $1/N$ (so that there is on average one jump by unit of time), and therefore it satisfies $\text{CDE}'(1/N, \infty)$. \square

Remark 6.2. We have shown that for the two-point space, the exponential curvature and the curvature are the same, and equal to 1. In [26], it is stated that the curvature is 2. The difference is because, since we enforced the normalization condition (1), the definitions of L in the two frameworks differ by a factor 2.

Acknowledgements

M. Fathi thanks Matthias Erbar, Ivan Gentil and Prasad Tetali for discussions on this topic. This work was partly done during a stay at the Hausdorff Research Institute for Mathematics in Bonn, whose support is gratefully acknowledged. This work was supported in part by NSF FRG Grant DMS-13-61185.

Y. Shu thanks his PhD advisors Nathaël Gozlan and Cyril Roberto for helpful advice and remarks.

We thank the referee for his valuable advice and comments.

References

- [1] Ambrosio, L., Gigli, N. and Savaré, G. (2015). Bakry–Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *Ann. Probab.* **43** 339–404. [MR3298475](#)

- [2] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C. and Scheffer, G. (2000). *Sur les Inégalités de Sobolev Logarithmiques. Panoramas et Synthèses [Panoramas and Syntheses]* **10**. Paris: Société Mathématique de France. [MR1845806](#)
- [3] Bakry, D. and Émery, M. (1985). Diffusions hypercontractives. In *Séminaire de Probabilités, XIX, 1983/84. Lecture Notes in Math.* **1123** 177–206. Berlin: Springer. [MR0889476](#)
- [4] Bakry, D., Gentil, I. and Ledoux, M. (2014). *Analysis and Geometry of Markov Diffusion Operators. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Cham: Springer. [MR3155209](#)
- [5] Bauer, F., Horn, P., Lin, Y., Lippner, G., Mangoubi, D. and Yau, S.-T. (2013). Li-Yau inequality on graphs. *J. Differential Geom.* **99** 359–405.
- [6] Bauer, F., Lin, Y., Liu, S. and Yau, S.-T. (2014). Volume doubling, Poincaré inequality and Gaussian heat kernel estimate for nonnegative curvature graphs.
- [7] Bobkov, S.G. and Tetali, P. (2006). Modified logarithmic Sobolev inequalities in discrete settings. *J. Theoret. Probab.* **19** 289–336. [MR2283379](#)
- [8] Bonciocat, A.-I. and Sturm, K.-T. (2009). Mass transportation and rough curvature bounds for discrete spaces. *J. Funct. Anal.* **256** 2944–2966. [MR2502429](#)
- [9] Buble, R. and Dyer, M.E. (1997). Path coupling: A technique for proving rapid mixing in Markov chains. In *38th Annual Symposium on Foundations of Computer Science, FOCS97* 223–231. Florida, USA: Miami Beach.
- [10] Chung, F., Lin, Y. and Yau, S.-T. (2014). Harnack inequalities for graphs with non-negative Ricci curvature. *J. Math. Anal. Appl.* **415** 25–32. [MR3173151](#)
- [11] Costello, K., Gozlan, N., Melbourne, J., Perkins, W., Roberto, C., Samson, P.-M., Shu, Y. and Tetali, P. Discrete Ricci curvature and functional inequalities, on graphs. In preparation.
- [12] Dobrušin, R.L. (1970). Definition of a system of random variables by means of conditional distributions. *Teor. Veroyatnost. i Primenen.* **15** 469–497. [MR0298716](#)
- [13] Eldan, R., Lee, J. and Lehec, J. (2016). Transport-entropy inequalities and curvature in discrete-space Markov chains. Preprint. Available at [arXiv:1604.06859](#).
- [14] Erbar, M. and Maas, J. (2012). Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206** 997–1038. [MR2989449](#)
- [15] Gozlan, N. (2009). A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.* **37** 2480–2498. [MR2573565](#)
- [16] Gozlan, N., Roberto, C. and Samson, P.-M. (2014). Hamilton–Jacobi equations on metric spaces and transport entropy inequalities. *Rev. Mat. Iberoam.* **30** 133–163. [MR3186934](#)
- [17] Gozlan, N., Roberto, C., Samson, P.-M. and Tetali, P. (2014). Displacement convexity of entropy and related inequalities on graphs. *Probab. Theory Related Fields* **160** 47–94. [MR3256809](#)
- [18] Gozlan, N., Roberto, C., Samson, P.-M. and Tetali, P. (2015). Kantorovich duality for general costs and applications. Preprint. Available at [arXiv:1412.7480](#).
- [19] Guillin, A. and Joulin, A. (2013). Measure concentration through non-Lipschitz observables and functional inequalities. *Electron. J. Probab.* **18** no. 65, 26. [MR3078024](#)
- [20] Guillin, A., Léonard, C., Wang, F.-Y. and Wu, L. (2009). Transportation-information inequalities for Markov processes (II): Relations with other functional inequalities. Unpublished note.
- [21] Guillin, A., Léonard, C., Wu, L. and Yao, N. (2009). Transportation-information inequalities for Markov processes. *Probab. Theory Related Fields* **144** 669–695. [MR2496446](#)
- [22] Johnson, O. (2015). A discrete log-Sobolev inequality under a Bakry–Émery type condition. *Ann. I.H.P. Probab. Stat.* To appear.
- [23] Jost, J. and Liu, S. (2014). Ollivier’s Ricci curvature, local clustering and curvature-dimension inequalities on graphs. *Discrete Comput. Geom.* **51** 300–322. [MR3164168](#)

- [24] Joulin, A. (2007). Poisson-type deviation inequalities for curved continuous-time Markov chains. *Bernoulli* **13** 782–798. [MR2348750](#)
- [25] Joulin, A. and Ollivier, Y. (2010). Curvature, concentration and error estimates for Markov chain Monte Carlo. *Ann. Probab.* **38** 2418–2442. [MR2683634](#)
- [26] Klartag, B., Kozma, G., Ralli, P. and Tetali, P. (2016). Discrete curvature and Abelian groups. *Canad. J. Math.* **68** 655–674. [MR3492631](#)
- [27] Ledoux, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Providence, RI: Amer. Math. Soc. [MR1849347](#)
- [28] Léonard, C. (2013). On the convexity of the entropy along entropic interpolations. Preprint.
- [29] Lin, Y. and Yau, S.-T. (2010). Ricci curvature and eigenvalue estimate on locally finite graphs. *Math. Res. Lett.* **17** 343–356. [MR2644381](#)
- [30] Lott, J. and Villani, C. (2009). Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)* **169** 903–991. [MR2480619](#)
- [31] Luczak, M.J. (2008). Concentration of measure and mixing for Markov chains. In *Fifth Colloquium on Mathematics and Computer Science. Discrete Math. Theor. Comput. Sci. Proc., AI 95–120*. Nancy: Assoc. Discrete Math. Theor. Comput. Sci. [MR2508781](#)
- [32] Maas, J. (2011). Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261** 2250–2292. [MR2824578](#)
- [33] Marton, K. (1996). Bounding \bar{d} -distance by informational divergence: A method to prove measure concentration. *Ann. Probab.* **24** 857–866. [MR1404531](#)
- [34] Mielke, A. (2013). Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differential Equations* **48** 1–31. [MR3090532](#)
- [35] Munch, F. (2015). Remarks on curvature dimension conditions on graphs. Preprint.
- [36] Oliveira, R.I. (2009). On the convergence to equilibrium of Kac’s random walk on matrices. *Ann. Appl. Probab.* **19** 1200–1231. [MR2537204](#)
- [37] Ollivier, Y. (2009). Ricci curvature of Markov chains on metric spaces. *J. Funct. Anal.* **256** 810–864. [MR2484937](#)
- [38] Ollivier, Y. (2010). A survey of Ricci curvature for metric spaces and Markov chains. In *Probabilistic Approach to Geometry. Adv. Stud. Pure Math.* **57** 343–381. Tokyo: Math. Soc. Japan. [MR2648269](#)
- [39] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. [MR1760620](#)
- [40] Sammer, M.D. (2005). Aspects of mass transportation in discrete concentration inequalities. Ph.D. thesis, Georgia Institute of Technology.
- [41] Schmuckenschläger, M. (1999). Curvature of nonlocal Markov generators. In *Convex Geometric Analysis (Berkeley, CA, 1996). Math. Sci. Res. Inst. Publ.* **34** 189–197. Cambridge: Cambridge Univ. Press. [MR1665591](#)
- [42] Shu, Y. (2014). Hamilton–Jacobi equations on graphs and applications. Preprint hal-01097744.
- [43] Sturm, K.-T. (2006). On the geometry of metric measure spaces. I. *Acta Math.* **196** 65–131. [MR2237206](#)
- [44] Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* **6** 587–600. [MR1392331](#)
- [45] Villani, C. (2009). *Optimal Transport, Old and New. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Berlin: Springer. [MR2459454](#)

Received October 2015 and revised July 2016