Available online at www.sciencedirect.com

ScienceDirect

Journal of Functional Analysis 256 (2009) 2944–2966

JOURNAL OF Functional Analysis

www.elsevier.com/locate/jfa

Mass transportation and rough curvature bounds for discrete spaces

Anca-Iuliana Bonciocat [a], Karl-Theodor Sturm [b,*]

[a] *Institute of Mathematics "S. Stoilow" of the Romanian Academy, P.O. Box 1-764, 014700 Bucharest, Romania*
[b] *Institute for Applied Mathematics, University of Bonn, Poppelsdorfer Allee 82/1, 53115 Bonn, Germany*

Received 2 August 2008; accepted 16 January 2009

Available online 20 February 2009

Communicated by C. Villani

Abstract

We introduce and study rough (approximate) lower curvature bounds for discrete spaces and for graphs. This notion agrees with the one introduced in [J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. 169 (2009), in press] and [K.T. Sturm, On the geometry of metric measure spaces. I, Acta Math. 196 (2006) 65–131], in the sense that the metric measure space which is approximated by a sequence of discrete spaces with rough curvature $\geqslant K$ will have curvature $\geqslant K$ in the sense of [J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. 169 (2009), in press; K.T. Sturm, On the geometry of metric measure spaces. I, Acta Math. 196 (2006) 65–131]. Moreover, in the converse direction, discretizations of metric measure spaces with curvature $\geqslant K$ will have rough curvature $\geqslant K$. We apply our results to concrete examples of homogeneous planar graphs.

Keywords: Optimal transport; Ricci curvature; GH-limits; Graphs; Concentration of measure

1. Introduction

We develop a notion of rough curvature bounds for discrete spaces, based on the concept of optimal mass transportation. These rough curvature bounds will depend on a real parameter $h > 0$, which should be considered as a natural length scale of the underlying discrete space or as

* Corresponding author.
E-mail addresses: Anca.Bonciocat@imar.ro (A.-I. Bonciocat), sturm@uni-bonn.de (K.-T. Sturm).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2009.01.029

the scale on which we have to look at the space. For a metric graph, for instance, this parameter equals the maximal length of its edges (times some constant).

The approach presented here will follow the one from [12], where the second author introduced a notion of lower curvature bounds for metric measure spaces, which is based on the concept of mass transportation. A closely related theory has been developed independently by J. Lott and C. Villani in [8], see also [15]. Both these approaches required the Wasserstein space of probability measures (and thus in turn the underlying space) to be a geodesic space. Therefore, in the original form they will not apply to discrete spaces. Moreover, if we consider a graph, more precisely the union of the edges of a graph, as a metric space it will have no lower curvature bound in the sense of [12], since the vertices will be branch points of geodesics which destroy the K-convexity of the entropy. The modification to be presented here overcomes this difficulty in the following way: mass transportation and convexity properties of the relative entropy will be studied along h-geodesics. For instance, instead of midpoints of a given pair of points x_0, x_1 we look at h-midpoints which are points y with $\mathsf{d}(x_0, y) \leqslant \frac{1}{2}\mathsf{d}(x_0, x_1) + h$ and $\mathsf{d}(x_1, y) \leqslant \frac{1}{2}\mathsf{d}(x_0, x_1) + h$.

Our first main result (Theorem 3.10) states that an arbitrary metric measure space (M, d, m) has curvature $\geqslant K$ (in the sense of [12]) provided it can be approximated by a sequence (M_h, d_h, m_h) of ('discrete') metric measure spaces with $h\text{-}\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K_h$ with $K_h \to K$ as $h \to 0$. That is, this result allows to pass from discrete spaces to continuous limit spaces.

Our second main result (Theorem 4.1) states that curvature bounds will also be preserved under the converse procedure: Given any metric space (M, d, m) with curvature $\geqslant K$ and any $h > 0$ we define standard discretizations (M_h, d, m_h) of (M, d, m) with $\mathbb{D}^2((M_h, \mathsf{d}, m_h), (M, \mathsf{d}, m)) \to 0$ as $h \to 0$ and with $h\text{-}\mathbb{C}\mathrm{urv}(M_h, \mathsf{d}_h, m_h) \geqslant K$.

Further, we apply our results to concrete examples. We prove (Theorem 5.3) that every homogeneous planar graph has h-curvature $\geqslant K$ where K is given in terms of the degree, the dual degree and the edge length. To be more precise, both the set $M = V$ of vertices, equipped with the counting measure, as well as the union $M = \bigcup_{e \in E} e$ of edges equipped with one-dimensional Lebesgue measure will be metric measure spaces with h-curvature $\geqslant K$, where the metric is the one induced by the Riemannian distance of the 2-dimensional Riemannian manifold whose discretization will be our given graph. Our notion of h-curvature yields the precise value for K if we consider discretizations of hyperbolic spaces.

In the final section we show that positive rough curvature bound implies a perturbed transportation cost inequality, weaker than what is usually called the Talagrand inequality. However, it still implies concentration of the reference measure m and exponential integrability of the Lipschitz functions with respect to m.

An independent, alternative approach to generalized Ricci curvature bounds for discrete spaces—again based on optimal transportation—was presented by Yann Ollivier [10], see Remark 6.4.

2. Preliminaries

Throughout this paper, a *metric measure space* will always be a triple (M, d, m) where (M, d) is a complete separable metric space and m is a measure on M (equipped with its Borel σ-algebra $\mathcal{B}(M)$) which is locally finite in the sense that $m(B_r(x)) < \infty$ for all $x \in M$ and all sufficiently small $r > 0$. We say that the metric measure space (M, d, m) is *normalized* if $m(M) = 1$.

Two metric measure spaces (M, d, m) and (M', d', m') are called *isomorphic* iff there exists an isometry $\psi : M_0 \to M_0'$ between the supports $M_0 := \mathrm{supp}[m] \subset M$ and $M_0' := \mathrm{supp}[m'] \subset M'$

such that $\psi_* m = m'$. The *diameter* of a metric measure space (M, d, m) will be the diameter of the metric space $(\mathrm{supp}[m], d)$.

We shall use the notion of *L_2-transportation distance* $\mathbb{D}$ for two metric measure spaces (M, d, m) and (M', d', m'), as defined in [12]:

$$\mathbb{D}\big((M, \mathsf{d}, m), (M', \mathsf{d}', m')\big) = \inf\Bigg(\int\limits_{M \sqcup M'} \hat{\mathsf{d}}^2(x, y)\, dq(x, y) \Bigg)^{1/2},$$

where $\hat{\mathsf{d}}$ ranges over all couplings of d and d' and q ranges over all couplings of m and m'. Here a measure q on the product space $M \times M'$ is a *coupling of m and m'* if $q(A \times M') = m(A)$ and $q(M \times A') = m'(A')$ for all measurable $A \subset M$, $A' \subset M'$; a pseudo-metric $\hat{\mathsf{d}}$ on the disjoint union $M \sqcup M'$ is a *coupling of* d *and* d' if $\hat{\mathsf{d}}(x, y) = \mathsf{d}(x, y)$ and $\hat{\mathsf{d}}(x', y') = \mathsf{d}'(x', y')$ for all $x, y \in \mathrm{supp}[m] \subset M$ and all $x', y' \in \mathrm{supp}[m'] \subset M'$.

The L_2-transportation distance $\mathbb{D}$ defines a complete and separable length metric on the family of all isomorphism classes of normalized metric measure spaces (M, d, m) for which $\int_M \mathsf{d}^2(z, x) dm(x) < \infty$ for some (hence all) $z \in M$. The notion of $\mathbb{D}$-convergence is closely related to the one of measured Gromov–Hausdorff convergence introduced in [4].

Recall that a sequence of compact normalized metric measure spaces $\{(M_n, \mathsf{d}_n, m_n)\}_{n \in \mathbb{N}}$ converges in the sense of *measured Gromov–Hausdorff convergence* (briefly, mGH-converges) to a compact normalized metric measure space (M, d, m) iff there exist a sequence of numbers $\epsilon_n \searrow 0$ and a sequence of measurable maps $f_n : M_n \to M$ such that for all $x, y \in M_n$, $|\mathsf{d}(f_n(x), f_n(y)) - \mathsf{d}_n(x, y)| \leqslant \epsilon_n$, for any $x \in M$ there exists $y \in M_n$ with $\mathsf{d}(f_n(y), x) \leqslant \epsilon_n$ and such that $(f_n)_* m_n \to m$ weakly on M for $n \to \infty$. According to Lemma 3.17 in [12], any mGH-convergent sequence of normalized metric measure spaces is also $\mathbb{D}$-convergent; for any sequence of normalized compact metric measure spaces with full supports and with uniform bounds for the doubling constants and for the diameters the notion of mGH-convergence is equivalent to the one of $\mathbb{D}$-convergence.

It is easy to see that $\mathbb{D}((M, \mathsf{d}, m), (M', \mathsf{d}', m')) = \inf \hat{\mathsf{d}}_W(\psi_* m, \psi'_* m')$ where the inf is taken over all metric spaces $(\hat{M}, \hat{\mathsf{d}})$ with isometric embeddings $\psi : M_0 \hookrightarrow \hat{M}$, $\psi' : M'_0 \hookrightarrow \hat{M}$ of the supports M_0 and M'_0 of m and m', respectively, and where $\hat{\mathsf{d}}_W$ denotes the L_2-Wasserstein distance derived from the metric $\hat{\mathsf{d}}$. Recall that for any metric space (M, d) the *L_2-Wasserstein distance* between two measures μ and ν on M is defined as

$$\mathsf{d}_W(\mu, \nu) = \inf\Bigg\{ \Bigg(\int\limits_{M \times M} \mathsf{d}^2(x, y)\, dq(x, y) \Bigg)^{1/2} : q \text{ is a coupling of } \mu \text{ and } \nu \Bigg\},$$

with the convention $\inf \emptyset = \infty$. For further details about the Wasserstein distance see the monograph [14]. We denote by $\mathcal{P}_2(M, \mathsf{d})$ the space of all probability measures ν which have finite second moments $\int_M \mathsf{d}^2(o, x)\, d\nu(x) < \infty$ for some (hence all) $o \in M$.

For a given metric measure space (M, d, m) we put $\mathcal{P}_2(M, \mathsf{d}, m)$ the space of all probability measures $\nu \in \mathcal{P}_2(M, \mathsf{d})$ which are absolutely continuous w.r.t. m. If $\nu = \rho \cdot m \in \mathcal{P}_2(M, \mathsf{d}, m)$ we consider the *relative entropy* of ν with respect to m defined by $\mathrm{Ent}(\nu | m) := \lim_{\epsilon \searrow 0} \int_{\{\rho > \epsilon\}} \rho \log \rho\, dm$. We denote by $\mathcal{P}_2^*(M, \mathsf{d}, m)$ the subspace of measures $\nu \in \mathcal{P}_2(M, \mathsf{d}, m)$ of finite entropy $\mathrm{Ent}(\nu \mid m) < \infty$.

We recall here the definitions of the lower curvature bounds for metric measure spaces introduced in [12]:

(i) A metric measure space (M, d, m) has *curvature* $\geqslant K$ for some number $K \in \mathbb{R}$ iff the relative entropy $\mathrm{Ent}(\cdot \mid m)$ is weakly K-convex on $\mathcal{P}_2^*(M, \mathsf{d}, m)$ in the sense that for each pair $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ there exists a geodesic $\Gamma : [0, 1] \to \mathcal{P}_2^*(M, \mathsf{d}, m)$ connecting ν_0 and ν_1 with

$$\begin{aligned} \mathrm{Ent}\big(\Gamma(t) \bigm| m\big) \leqslant{}& (1-t)\,\mathrm{Ent}\big(\Gamma(0) \bigm| m\big) + t\,\mathrm{Ent}\big(\Gamma(1) \bigm| m\big) \\ &- \frac{K}{2} t(1-t) \mathsf{d}_W^2\big(\Gamma(0), \Gamma(1)\big) \end{aligned} \tag{2.1}$$

for all $t \in [0, 1]$.

(ii) The metric measure space (M, d, m) has *curvature* $\geqslant K$ *in the lax sense* iff for each $\epsilon > 0$ and for each pair $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ there exists an ϵ-midpoint $\eta \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ of ν_0 and ν_1 with

$$\mathrm{Ent}(\eta \mid m) \leqslant \frac{1}{2} \mathrm{Ent}(\nu_0 \mid m) + \frac{1}{2} \mathrm{Ent}(\nu_1 \mid m) - \frac{K}{8} \mathsf{d}_W^2(\nu_0, \nu_1) + \epsilon. \tag{2.2}$$

Briefly, we shall write $\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K$, respectively $\mathbb{C}\mathrm{urv}_{lax}(M, \mathsf{d}, m) \geqslant K$.

Recall that in a given metric space (M, d) a point y is an ϵ-*midpoint* of x_0 and x_1 if $\mathsf{d}(x_i, y) \leqslant \frac{1}{2}\mathsf{d}(x_0, x_1) + \epsilon$ for each $i = 0, 1$. We call y *midpoint* of x_0 and x_1 if $\mathsf{d}(x_i, y) \leqslant \frac{1}{2}\mathsf{d}(x_0, x_1)$ for $i = 0, 1$.

3. Rough curvature bounds for metric measure spaces

In order to adapt the notion of curvature bound to other spaces then geodesic without branching we shall refer in this paper to a larger class of metric spaces:

Definition 3.1. Let $h > 0$ be given. We say that a metric space (M, d) is *h-rough geodesic* iff for each pair of points $x_0, x_1 \in M$ and each $t \in [0, 1]$ there exists a point $x_t \in M$ satisfying

$$\mathsf{d}(x_0, x_t) \leqslant t\mathsf{d}(x_0, x_1) + h, \qquad \mathsf{d}(x_t, x_1) \leqslant (1-t)\mathsf{d}(x_0, x_1) + h. \tag{3.1}$$

The point x_t will be referred to as the h-rough t-approximate point between x_0 and x_1. The h-rough $1/2$-approximate point is actually the h-midpoint of x_0 and x_1.

Example 3.2.

(i) Any nonempty set X with the discrete metric $\mathsf{d}(x, y) = 0$ for $x = y$ and 1 for $x \neq y$ is h-rough geodesic for any $h \geqslant 1/2$. In this case, any point is an h-midpoint of any pair of distinct points.

(ii) If $\epsilon > 0$ then the space $(\mathbb{R}^n, \mathsf{d})$ with the metric $\mathsf{d}(x, y) = |x - y| \wedge \epsilon$ is h-rough geodesic for $h \geqslant \epsilon/2$ (here $|\cdot|$ is the euclidian metric).

(iii) For $\epsilon > 0$ the space $(\mathbb{R}^n, \mathsf{d})$ with the metric $\mathsf{d}(x, y) = \sqrt{\epsilon|x - y| + |x - y|^2}$ is h-rough geodesic for each $h \geqslant \epsilon/4$.

The above examples are somewhat pathological. We actually have in mind the more friendly examples of discrete spaces and some geodesic spaces with branch points, e.g. graphs, that do not have curvature bounds as defined in [12].

For a discrete h-rough geodesic metric space (M, d) one should think of h as a discretization size or "resolution" of M. In an h-geodesic space a pair of points x and y is not necessarily connected by a geodesic but by a chain of points $x = x_0, x_1, \dots, x_n = y$ having intermediate distance less then $h/2$.

In the sequel we will use two types of perturbations of the Wasserstein distance, defined as follows:

Definition 3.3. Let (M, d) be a metric space. For each $h > 0$ and any pair of measures $\nu_0, \nu_1 \in \mathcal{P}_2(M, \mathsf{d})$ put

$$\mathsf{d}_W^{\pm h}(\nu_0, \nu_1) := \inf\left\{ \left(\int \left[\left(\mathsf{d}(x_0, x_1) \mp h \right)_+ \right]^2 dq(x_0, x_1) \right)^{1/2} : q \text{ coupling of } \nu_0 \text{ and } \nu_1 \right\}, \tag{3.2}$$

where $(\cdot)_+$ denotes the positive part.

Remark 3.4. According to Theorem 4.1 from [15] there exists a coupling for which the infimum in (3.2) is attaint. We will call it $+h$-optimal coupling (resp. $-h$-optimal coupling) of ν_0 and ν_1.

The two perturbations d_W^{+h} and d_W^{-h} are related to the Wasserstein distance d_W in the following way:

Lemma 3.5. *For any $h > 0$ we have*

(i) $\mathsf{d}_W^{+h} \leqslant \mathsf{d}_W \leqslant \mathsf{d}_W^{+h} + h$;
(ii) $\mathsf{d}_W \leqslant \mathsf{d}_W^{-h} \leqslant \mathsf{d}_W + h$.

Proof. (i) Let ν_0 and ν_1 be two probabilities in (M, d) and consider q an optimal coupling and q_{+h} a $+h$-optimal coupling of them. Then

$$\begin{aligned} \mathsf{d}_W^{+h}(\nu_0, \nu_1) &= \left(\int \left[\left(\mathsf{d}(x_0, x_1) - h \right)_+ \right]^2 dq_{+h}(x_0, x_1) \right)^{1/2} \\ &\leqslant \left(\int \left[\left(\mathsf{d}(x_0, x_1) - h \right)_+ \right]^2 dq(x_0, x_1) \right)^{1/2} \\ &\leqslant \left(\int \mathsf{d}(x_0, x_1)^2 dq(x_0, x_1) \right)^{1/2} = \mathsf{d}_W(\nu_0, \nu_1) \end{aligned}$$

and

$$\begin{aligned} \mathsf{d}_W(\nu_0, \nu_1) &= \left(\int \mathsf{d}(x_0, x_1)^2 \, dq(x_0, x_1) \right)^{1/2} \leqslant \left(\int \mathsf{d}(x_0, x_1)^2 \, dq_{+h}(x_0, x_1) \right)^{1/2} \\ &\leqslant \left(\int \left[\left(\mathsf{d}(x_0, x_1) - h \right)_+ + h \right]^2 dq_{+h}(x_0, x_1) \right)^{1/2} \leqslant \mathsf{d}_W^{+h}(\nu_0, \nu_1) + h. \end{aligned}$$

(ii) Similar to (i). $\square$

With an elementary proof we have also a monotonicity property of $\mathsf{d}_W^{\pm h}$ in h:

Lemma 3.6. *Let $0 < h_1 < h_2$ be arbitrarily given. Then for each pair of probabilities ν_0 and ν_1*

(i) $\mathsf{d}_W^{-h_1}(\nu_0, \nu_1) < \mathsf{d}_W^{-h_2}(\nu_0, \nu_1)$;
(ii) $\mathsf{d}_W^{+h_1}(\nu_0, \nu_1) \geqslant \mathsf{d}_W^{+h_2}(\nu_0, \nu_1)$ *and the inequality is strict if and only if* $\mathsf{d}_W^{+h_1}(\nu_0, \nu_1) > 0$.

We introduce now the notion of rough lower curvature bound:

Definition 3.7. We say that a metric measure space (M, d, m) has *h-rough curvature* $\geqslant K$ for some numbers $h > 0$ and $K \in \mathbb{R}$ iff for each pair $\nu_0, \nu_1 \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ and for any $t \in [0, 1]$ there exists an h-rough t-approximate point $\eta_t \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ between ν_0 and ν_1 satisfying

$$\mathrm{Ent}(\eta_t \mid m) \leqslant (1-t)\,\mathrm{Ent}(\nu_0 \mid m) + t\,\mathrm{Ent}(\nu_1 \mid m) - \frac{K}{2} t(1-t)\mathsf{d}_W^{\pm h}(\nu_0, \nu_1)^2, \tag{3.3}$$

where the sign in $\mathsf{d}_W^{\pm h}(\nu_0, \nu_1)$ is chosen '$+$' if $K > 0$ and '$-$' if $K < 0$. Briefly, we write in this case $h\text{-}\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K$.

Remark 3.8. We could also choose two parameters in the above definition, h for the approximate midpoint and ϵ for the inequality (3.3). Having two parameters instead of one is not essentially useful for further results. One can always think of $h \vee \epsilon$ in the definition of rough curvature bound, which is an approximate notion.

Remark 3.9. (i) If (M, d, m) and (M', d', m') are two isomorphic metric measure spaces and $K \in \mathbb{R}$ then $h\text{-}\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K$ if and only if $h\text{-}\mathbb{C}\mathrm{urv}(M', \mathsf{d}', m') \geqslant K$.

(ii) If (M, d, m) is a metric measure space and $\alpha, \beta > 0$ then $h\text{-}\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K$ if and only if $\alpha h\text{-}\mathbb{C}\mathrm{urv}(M, \alpha\mathsf{d}, \beta m) \geqslant \frac{K}{\alpha^2}$, because $\mathrm{Ent}(\nu \mid \beta m) = \mathrm{Ent}(\nu \mid m) - \log\beta$, $(\alpha \cdot \mathsf{d})_W^{\pm h}(\nu_0, \nu_1) = \alpha \cdot \mathsf{d}_W^{\pm h}(\nu_0, \nu_1)$ and for $t \in [0, 1]$ η_t is h-rough t-approximate point between μ, ν with respect to d_W if and only if η_t is αh-rough t-approximate point between μ, ν with respect to $(\alpha\mathsf{d})_W$.

Theorem 3.10. *Let (M, d, m) be a normalized metric measure space and $\{(M_h, \mathsf{d}_h, m_h)\}_{h>0}$ a family of normalized metric measure spaces with uniformly bounded diameter and with h-$\mathbb{C}\mathrm{urv}(M_h, \mathsf{d}_h, m_h) \geqslant K_h$ for $K_h \to K$ as $h \to 0$. If*

$$(M_h, \mathsf{d}_h, m_h) \xrightarrow{\mathbb{D}} (M, \mathsf{d}, m)$$

as $h \to 0$ then

$$\mathbb{C}\mathrm{urv}_{lax}(M, \mathsf{d}, m) \geqslant K.$$

If in addition M is compact then

$$\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K.$$

Proof. Let $\{(M_h, \mathsf{d}_h, m_h)\}_{h>0}$ be a family of normalized discrete metric measure spaces. Assume that $(M_h, \mathsf{d}_h, m_h) \xrightarrow{\mathbb{D}} (M, \mathsf{d}, m)$ as $h \to 0$ and $\sup_{h>0} \mathrm{diam}(M_h, \mathsf{d}_h, m_h)$,

$\operatorname{diam}(M, \mathsf{d}, m) \leqslant \Delta$ for some $\Delta \in \mathbb{R}$. Now let $\epsilon > 0$ and $\nu_0 = \rho_0 m$, $\nu_1 = \rho_1 m \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ be given. Choose R with

$$\sup_{i=0,1} \operatorname{Ent}(\nu_i \mid m) + \frac{|K|}{8}\Delta^2 + \frac{\epsilon}{8}\big[\Delta^2 + 3|K|(2\Delta + 3\epsilon)\big] \leqslant R. \tag{3.4}$$

We have to deduce the existence of an ϵ-midpoint η which satisfies inequality (2.2). Choose $0 < h < \epsilon$ with $|K_h - K| < \epsilon$ and

$$\mathbb{D}\big((M_h, \mathsf{d}_h, m_h), (M, \mathsf{d}, m)\big) \leqslant \exp\bigg(-\frac{2 + 4\Delta^2 R}{\epsilon^2}\bigg). \tag{3.5}$$

Like in Section 4.5 in [12], one can define the canonical maps $Q'_h : \mathcal{P}_2(M, \mathsf{d}, m) \to \mathcal{P}_2(M_h, \mathsf{d}_h, m_h)$, and $Q_h : \mathcal{P}_2(M_h, \mathsf{d}_h, m_h) \to \mathcal{P}_2(M, \mathsf{d}, m)$ as follows.

We consider q_h a coupling of m and m_h and $\hat{\mathsf{d}}_h$ a coupling of d and d_h such that

$$\int \hat{\mathsf{d}}_h^2(x, y)\, dq_h(x, y) \leqslant 2\mathbb{D}^2\big((M, \mathsf{d}, m), (M_h, \mathsf{d}_h, m_h)\big).$$

Let Q'_h and Q_h be the disintegrations of q_h w.r.t. m_h and m, resp., that is $dq_h(x, y) = Q'_h(y, dx)\, dm_h(y) = Q_h(x, dy)\, dm(x)$ and let $\hat{\Delta}$ denote the m-essential supremum of the map

$$x \mapsto \bigg[\int_{M_h} \hat{\mathsf{d}}_h^2(x, y)\, Q_h(x, dy)\bigg]^{1/2}.$$

In our case $\hat{\Delta} \leqslant 2\Delta$.

For $\nu = \rho m \in \mathcal{P}_2(M, \mathsf{d}, m)$ define $Q'_h(\nu) \in \mathcal{P}_2(M_h, \mathsf{d}_h, m_h)$ by $Q'_h(\nu) := \rho_h m_h$ where

$$\rho_h(y) := \int_M \rho(x)\, Q'_h(y, dx).$$

The map Q_h is defined similarly. Lemma 4.19 from [12] gives the following estimates:

$$\operatorname{Ent}\big(Q'_h(\nu) \,\big|\, m_h\big) \leqslant \operatorname{Ent}(\nu \mid m) \quad \text{for all } \nu = \rho m \tag{3.6}$$

and

$$\mathsf{d}_W^2\big(\nu, Q'_h(\nu)\big) \leqslant \frac{2 + \hat{\Delta}^2 \cdot \operatorname{Ent}(\nu \mid m)}{-\log \mathbb{D}((M, \mathsf{d}, m), (M_h, \mathsf{d}_h, m_h))}. \tag{3.7}$$

provided $\mathbb{D}((M, \mathsf{d}, m), (M_h, \mathsf{d}_h, m_h)) < 1$. Analogous estimates hold for Q_h.

For our given $\nu_0 = \rho_0 m$, $\nu_1 = \rho_1 m \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ put

$$\nu_{i,h} := Q'_h(\nu_i) = \rho_{i,h} m_h$$

with $\rho_{i,h}(y)=\int \rho_i(x)Q'_h(y,dx)$ for $i=0,1$ and let η_h be an h-midpoint of $\nu_{0,h}$ and $\nu_{1,h}$ such that

$$\mathrm{Ent}(\eta_h \mid m_h) \leqslant \frac{1}{2}\mathrm{Ent}(\nu_{0,h} \mid m_h) + \frac{1}{2}\mathrm{Ent}(\nu_{1,h} \mid m_h) - \frac{K_h}{8}\mathsf{d}_W^{\delta_h h}(\nu_{0,h},\nu_{1,h})^2, \tag{3.8}$$

where δ_h is the sign of K_h.

From (3.5)–(3.7) we conclude

$$\begin{aligned}\mathsf{d}_W^2(\nu_0,\nu_{0,h}) &\leqslant \frac{2+\hat{\Delta}^2\cdot \mathrm{Ent}(\nu_0 \mid m)}{-\log \mathbb{D}((M,\mathsf{d},m),(M_h,\mathsf{d}_h,m_h))}\\ &\leqslant \frac{2+4\Delta^2 R}{-\log \mathbb{D}((M,\mathsf{d},m),(M_h,\mathsf{d}_h,m_h))} \leqslant \epsilon^2\end{aligned}$$

and similarly $\mathsf{d}_W^2(\nu_1,\nu_{1,h}) \leqslant \epsilon^2$.

If $K<0$ we can suppose $K_h<0$ too. From Lemma 3.5(ii) we have

$$\mathsf{d}_W^{-h}(\nu_{0,h},\nu_{1,h})^2 \leqslant \big(\mathsf{d}_W(\nu_{0,h},\nu_{1,h})+h\big)^2 \leqslant \big(\mathsf{d}_W(\nu_0,\nu_1)+3\epsilon\big)^2 \leqslant \mathsf{d}_W(\nu_0,\nu_1)^2+6\epsilon\Delta+9\epsilon^2,$$

because $\mathsf{d}_W(\nu_0,\nu_1) \leqslant \Delta$.

For $K>0$ one can choose h small enough to ensure $K_h>0$. Then Lemma 3.5(i) implies

$$\mathsf{d}_W(\nu_0,\nu_1)^2 \leqslant \big(\mathsf{d}_W(\nu_{0,h},\nu_{1,h})+2\epsilon\big)^2 \leqslant \big(\mathsf{d}_W^{+h}(\nu_{0,h},\nu_{1,h})+3\epsilon\big)^2 \leqslant \mathsf{d}_W^{+h}(\nu_0,\nu_1)^2+6\epsilon\Delta+9\epsilon^2.$$

In both cases the estimates above combined with (3.6), (3.8) and the fact that we chose h with $-K_h<\epsilon-K$ will imply

$$\mathrm{Ent}(\eta_h \mid m_h) \leqslant \frac{1}{2}\mathrm{Ent}(\nu_0 \mid m) + \frac{1}{2}\mathrm{Ent}(\nu_1 \mid m) - \frac{K}{8}\mathsf{d}_W^2(\nu_0,\nu_1)+\epsilon' \tag{3.9}$$

with $\epsilon'=\epsilon[\Delta^2+3|K|(2\Delta+3\epsilon)]/8$.

The case $K=0$ follows by the calculations above, depending on the sign of K_h.

Finally, put

$$\eta = Q_h(\eta_h).$$

Then again by (3.5), the estimates given in Lemma 4.19 [12] for Q_h and by the previous estimate (3.9) for $\mathrm{Ent}(\eta_h \mid m_h)$ we deduce

$$\begin{aligned}\mathsf{d}_W^2(\eta_h,\eta) &\leqslant \frac{2+\hat{\Delta}^2\cdot \mathrm{Ent}(\eta_h \mid m_h)}{-\log \mathbb{D}((M,\mathsf{d},m),(M_h,\mathsf{d}_h,m_h))}\\ &\leqslant \frac{2+4\Delta^2 R}{-\log \mathbb{D}((M,\mathsf{d},m),(M_h,\mathsf{d}_h,m_h))} \leqslant \epsilon^2.\end{aligned}$$

For $i = 0, 1$ we have $\mathsf{d}_W(\eta, \nu_i) \leqslant 2\epsilon + \mathsf{d}_W(\eta_h, \nu_{i,h}) \leqslant 2\epsilon + \frac{1}{2}\mathsf{d}_W(\nu_{0,h}, \nu_{1,h}) + h \leqslant \frac{1}{2}\mathsf{d}_W(\nu_0, \nu_1) + 4\epsilon$. Hence,

$$\sup_{i=0,1} \mathsf{d}_W(\eta, \nu_i) \leqslant \frac{1}{2}\mathsf{d}_W(\nu_0, \nu_1) + 4\epsilon,$$

i.e. η is a (4ϵ)-midpoint of ν_0 and ν_1. Furthermore, by (3.6)

$$\begin{aligned}\mathrm{Ent}(\eta \mid m) &\leqslant \mathrm{Ent}(\eta_h \mid m_h)\\ &\leqslant \frac{1}{2}\mathrm{Ent}(\nu_0 \mid m) + \frac{1}{2}\mathrm{Ent}(\nu_1 \mid m) - \frac{K}{8}\mathsf{d}_W^2(\nu_0, \nu_1) + \epsilon'\end{aligned}$$

with ϵ' as above. This proves that $\mathbb{C}\mathrm{urv}_{lax}(M, \mathsf{d}, m) \geqslant K$. □

4. Discretizations of metric spaces

Let (M, d, m) be a given metric measure space. For $h > 0$ let M_h be a discrete subset of M, say $M_h = \{x_n\colon n \in \mathbb{N}\}$, with $M = \bigcup_{i=1}^{\infty} B_R(x_i)$, where $R = R(h) \searrow 0$ as $h \searrow 0$. If (M, d, m) has finite diameter then M_h might consist of a finite number of points. Choose $A_i \subset B_R(x_i)$ mutually disjoint with $x_i \in A_i$, $i = 1, 2, \ldots$, and $\bigcup_{i=1}^{\infty} A_i = M$ (e.g. one could choose a Voronoi tessellation) and consider the measure m_h on M_h given by $m_h(\{x_i\}) := m(A_i)$, $i = 1, 2, \ldots$. We call (M_h, d, m_h) a *discretization* of (M, d, m).

Theorem 4.1.

(i) *If $m(M) < \infty$ then $(M_h, \mathsf{d}, m_h) \xrightarrow{\mathbb{D}} (M, \mathsf{d}, m)$ as $h \to 0$.*
(ii) *If $\mathbb{C}\mathrm{urv}_{lax}(M, \mathsf{d}, m) \geqslant K$ with $K \neq 0$ then for each $h > 0$ and for each discretization (M_h, d, m_h) with $R(h) < h/4$ we have h-$\mathbb{C}\mathrm{urv}(M_h, \mathsf{d}, m_h) \geqslant K$.*
(iii) *If $\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K$ for some real number K then for each $h > 0$ and for each discretization (M_h, d, m_h) with $R(h) \leqslant h/4$ we have h-$\mathbb{C}\mathrm{urv}(M_h, \mathsf{d}, m_h) \geqslant K$.*

Proof. (i) The measure $q = \sum_{i=1}^{\infty}(m(A_i)\delta_{x_i}) \times (1_{A_i} m)$ is a coupling of m_h and m, so

$$\begin{aligned}\mathbb{D}^2\big((M_h, \mathsf{d}, m_h), (M, \mathsf{d}, m)\big) &\leqslant \int_{M_h \times M} \mathsf{d}^2(x, y)\, dq(x, y)\\ &= \sum_{i=1}^{\infty} m(A_i) \int_{A_i} \mathsf{d}^2(x_i, y)\, dm(y)\\ &\leqslant \left(\sum_{i=1}^{\infty} m(A_i)^2\right) R(h)^2 \leqslant R(h)^2 \left(\sum_{i=1}^{\infty} m(A_i)\right)^2\\ &= R(h)^2 m(M)^2 \to 0 \quad \text{as } h \to 0.\end{aligned}$$

(ii) Fix $h > 0$ and consider a discretization (M_h, d, m_h) of (M, d, m) with $R(h) < h/4$. Let $\nu_0^h, \nu_1^h \in \mathcal{P}_2^*(M_h, \mathsf{d}, m_h)$ be given; it is enough to make the proof for ν_0^h, ν_1^h with compact support.

Suppose then $\nu_i^h = (\sum_{j=1}^n \alpha_{i,j}^h 1_{\{x_j\}}) m_h$, $i = 1, 2$ (some of the $\alpha_{i,j}^h$ can be zero). We take also an arbitrary $t \in [0,1]$. Put $\nu_i := (\sum_{j=1}^n \alpha_{i,j}^h 1_{A_j}) m \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ for $i = 1, 2$. Choose $\epsilon > 0$ such that

$$4R(h) + \epsilon \leqslant h. \tag{4.1}$$

Since $\mathbb{C}\mathrm{urv}_{lax}(M, \mathsf{d}, m) \geqslant K$ for our given $t \in [0,1]$ there exists $\eta_t \in \mathcal{P}_2^*(M, \mathsf{d}, m)$ an ϵ-rough t-approximate point between ν_0 and ν_1 such that

$$\mathrm{Ent}(\eta_t \mid m) \leqslant (1-t)\,\mathrm{Ent}(\nu_0 \mid m) + t\,\mathrm{Ent}(\nu_1 \mid m) - \frac{K}{2} t(1-t)\mathsf{d}_W^2(\nu_0, \nu_1) + \epsilon. \tag{4.2}$$

We compute

$$\mathrm{Ent}(\nu_i \mid m) = \sum_{j=1}^n \int_{A_j} \alpha_{i,j}^h \log \alpha_{i,j}^h \, dm = \sum_{j=1}^n \alpha_{i,j}^h \log \alpha_{i,j}^h m_h(\{x_j\}) = \mathrm{Ent}(\nu_i^h \mid m_h), \tag{4.3}$$

for $i = 0, 1$. Denote $\eta_t^h(\{x_j\}) := \eta_t(A_j)$, $j = 1, 2, \ldots, n$. Suppose $\eta_t = \rho_t \cdot m$. From Jensen's inequality we get

$$\begin{aligned}
\mathrm{Ent}(\eta_t^h \mid m_h) &= \sum_{j=1}^\infty \frac{\int_{A_j} \rho_t \, dm}{m(A_j)} \log \frac{\int_{A_j} \rho_t \, dm}{m(A_j)} m_h(\{x_j\}) \\
&\leqslant \sum_{j=1}^\infty \left(\frac{1}{m(A_j)} \int_{A_j} \rho_t \log \rho_t \, dm \right) m_h(\{x_j\}) = \mathrm{Ent}(\eta_t \mid m),
\end{aligned}$$

which together with (4.2) and (4.3) implies

$$\mathrm{Ent}(\eta_t^h \mid m_h) \leqslant (1-t)\,\mathrm{Ent}(\nu_0^h \mid m_h) + t\,\mathrm{Ent}(\nu_1^h \mid m_h) - \frac{K}{2} t(1-t)\mathsf{d}_W^2(\nu_0, \nu_1) + \epsilon. \tag{4.4}$$

Firstly, we consider the case $K < 0$. Let q^h be a $-2R(h)$-optimal coupling of ν_0^h and ν_1^h. Then the formula

$$\hat{q} := \sum_{j,k=1}^n \left[q^h(\{(x_j, x_k)\}) \delta_{(x_j, x_k)} \times \frac{1_{A_j \times A_k}}{m(A_j) m(A_k)} (m \times m) \right]$$

defines a measure on $M_h \times M_h \times M \times M$ which has marginals ν_0^h, ν_1^h, ν_0 and ν_1. Moreover, the projection of $\hat{q}$ on the first two factors is equal to q^h. Therefore we have

$$\begin{aligned}
\mathsf{d}_W(\nu_0, \nu_1)^2 &\leqslant \int \mathsf{d}(x, y)^2 d\hat{q}(x^h, y^h, x, y) \\
&\leqslant \int \left[\mathsf{d}(x, x^h) + \mathsf{d}(x^h, y^h) + \mathsf{d}(y^h, y) \right]^2 d\hat{q}(x^h, y^h, x, y)
\end{aligned}$$

$$= \sum_{j,k=1}^{n} \frac{q^h(\{(x_j, x_k)\})}{m(A_j)m(A_k)} \int_{A_j \times A_k} \big[\mathsf{d}(x, x_j) + \mathsf{d}(x_j, x_k) + \mathsf{d}(x_k, y)\big]^2 \, dm(x)\, dm(y)$$

$$\leqslant \sum_{j,k=1}^{n} q^h\big(\{(x_j, x_k)\}\big)\big(\mathsf{d}(x_j, x_k) + 2R(h)\big)^2 = \mathsf{d}_W^{-2R(h)}(\nu_0^h, \nu_1^h)^2,$$

which together with (4.4) yields

$$\begin{aligned} \mathrm{Ent}\big(\eta_t^h \,\big|\, m_h\big) &\leqslant (1-t)\,\mathrm{Ent}\big(\nu_0^h \,\big|\, m_h\big) + t\,\mathrm{Ent}\big(\nu_1^h \,\big|\, m_h\big) \\ &\quad - \frac{K}{2} t(1-t)\mathsf{d}_W^{-2R(h)}\big(\nu_0^h, \nu_1^h\big)^2 + \epsilon. \end{aligned} \tag{4.5}$$

In the case $K > 0$ we start with an optimal coupling q of ν_0 and ν_1 and we show that the measure

$$\tilde{q}^h := \sum_{j,k=1}^{n} q(A_j \times A_k)\delta_{(x_j, x_k)}$$

is a coupling of ν_0^h and ν_1^h. Indeed, if $A \subset M_h$ then we have in turn

$$\begin{aligned} \sum_{j,k=1}^{n} q(A_j \times A_k)\delta_{(x_j,x_k)}(A \times M_h) &= \sum_{j,k=1}^{n} q(A_j \times A_k)\delta_{x_j}(A) = \sum_{j=1}^{n} q(A_j \times M)\delta_{x_j}(A) \\ &= \sum_{j=1}^{n} \nu_0(A_j)\delta_{x_j}(A) = \sum_{j=1}^{n} \nu_0^h\big(\{x_j\}\big)\delta_{x_j}(A) = \nu_0^h(A). \end{aligned}$$

Since for any $j, k = 1, 2, \ldots, n$ and for arbitrary $x \in A_j$ and $y \in A_k$ we have $(\mathsf{d}(x_j, x_k) - 2R(h))_+ \leqslant (\mathsf{d}(x_j, x_k) - \mathsf{d}(x, x_j) - \mathsf{d}(y, x_k))_+ \leqslant \mathsf{d}(x, y)$ one can estimate:

$$\begin{aligned} \mathsf{d}_W^{+2R(h)}\big(\nu_0^h, \nu_1^h\big)^2 &\leqslant \sum_{j,k=1}^{n} q(A_j \times A_k)\big[\big(\mathsf{d}(x_j, x_k) - 2R(h)\big)_+\big]^2 \\ &= \sum_{j,k=1}^{n} \int_{A_j \times A_k} \big[\big(\mathsf{d}(x_j, x_k) - 2R(h)\big)_+\big]^2 \, dq(x, y) \\ &\leqslant \sum_{j,k=1}^{n} \int_{A_j \times A_k} \big[\big(\mathsf{d}(x_j, x_k) - \mathsf{d}(x, x_j) - \mathsf{d}(y, x_k)\big)_+\big]^2 \, dq(x, y) \\ &\leqslant \sum_{j,k=1}^{n} \int_{A_j \times A_k} \mathsf{d}(x, y)^2 \, dq(x, y) = \int_{M \times M} \mathsf{d}(x, y)^2 \, dq(x, y) = \mathsf{d}_W(\nu_0, \nu_1)^2. \end{aligned}$$

Therefore from (4.4) we obtain

$$\begin{aligned}\mathrm{Ent}\big(\eta_t^h \bigm| m_h\big) &\leqslant (1-t)\,\mathrm{Ent}\big(\nu_0^h \bigm| m_h\big) + t\,\mathrm{Ent}\big(\nu_1^h \bigm| m_h\big)\\ &\quad - \frac{K}{2}t(1-t)\mathsf{d}_W^{+2R(h)}\big(\nu_0^h, \nu_1^h\big)^2 + \epsilon. \end{aligned} \tag{4.6}$$

For ϵ sufficiently small we can get

$$-\frac{K}{2}t(1-t)\mathsf{d}_W^{\pm 2R(h)}\big(\nu_0^h, \nu_1^h\big)^2 + \epsilon \leqslant -\frac{K}{2}t(1-t)\mathsf{d}_W^{\pm h}\big(\nu_0^h, \nu_1^h\big)^2 \tag{4.7}$$

and then (4.5), (4.6) yield

$$\mathrm{Ent}\big(\eta_t^h \bigm| m_h\big) \leqslant (1-t)\,\mathrm{Ent}\big(\nu_0^h \bigm| m_h\big) + t\,\mathrm{Ent}\big(\nu_1^h \bigm| m_h\big) - \frac{K}{2}t(1-t)\mathsf{d}_W^{\pm h}\big(\nu_0^h, \nu_1^h\big)^2, \tag{4.8}$$

depending on the sign of K. The inequality (4.7) fails only when $K > 0$ and $\mathsf{d}_W^{+h}(\nu_0^h, \nu_1^h) = 0$, but in this case $\mathsf{d}_W(\nu_0^h, \nu_1^h) \leqslant h$ and either $\eta = \nu_0^h$ or $\eta = \nu_1^h$ verifies directly the condition (3.3) from the definition of h-rough curvature bound for the discretization.

The measure $\pi = \sum_{j=1}^n (\eta_t^h(\{x_j\})\delta_{x_j} \times 1_{A_j}\eta_t)$ is a coupling of η_t^h and η_t, so

$$\mathsf{d}_W^2\big(\eta_t^h, \eta_t\big) \leqslant \int\limits_{M_h \times M} \mathsf{d}^2(x, y)\, d\pi(x, y) \leqslant R^2(h),$$

and similarly $\mathsf{d}_W^2(\nu_i^h, \nu_i) \leqslant R^2(h)$ for $i = 1, 2$. Because η_t is an ϵ-rough t-approximate point between ν_0 and ν_1 we deduce

$$\begin{aligned}\mathsf{d}_W\big(\eta_t^h, \nu_0^h\big) &\leqslant \mathsf{d}_W(\eta_t, \nu_0) + 2R(h) \leqslant t\mathsf{d}_W(\nu_0, \nu_1) + 2R(h) + \epsilon\\ &\leqslant t\mathsf{d}_W\big(\nu_0^h, \nu_1^h\big) + 2R(h)(1+t) + \epsilon\end{aligned}$$

and by a similar argument

$$\mathsf{d}_W\big(\eta_t^h, \nu_1^h\big) \leqslant (1-t)\mathsf{d}_W\big(\nu_0^h, \nu_1^h\big) + 2R(h)(2-t) + \epsilon.$$

From (4.1) we conclude that η^h is an h-rough t-approximate point between ν_0^h and ν_1^h, which together with (4.8) proves that $h\text{-}\mathbb{C}\mathrm{urv}(M_h, \mathsf{d}, m_h) \geqslant K$.

(iii) follows the same lines as (ii). □

Example 4.2.

- (i) If we consider on $\mathbb{Z}^n$ the metric d_1 coming from the norm $|\cdot|_1$ in $\mathbb{R}^n$ defined by $|x|_1 = \sum_{i=1}^n |x_i|$ and with the measure $\overline{m}_n = \sum_{x \in \mathbb{Z}^n} \delta_x$, then $h\text{-}\mathbb{C}\mathrm{urv}(\mathbb{Z}^n, \mathsf{d}_1, \overline{m}_n) \geqslant 0$ for any $h \geqslant 2n$.
- (ii) The n-dimensional grid $\mathbb{E}^n$ having $\mathbb{Z}^n$ as set of vertices, equipped with the graph distance and with the measure m_n which is the 1-dimensional Lebesgue measure on the edges, has $h\text{-}\mathbb{C}\mathrm{urv}(\mathbb{E}^n, \mathsf{d}_1, m_n) \geqslant 0$ for any $h \geqslant 2(n+1)$.

Proof. We use the following result:

Lemma 4.3. *(See [15].) Any finite dimensional Banach space equipped with the Lebesgue measure has curvature* $\geqslant 0$.

We tile the space $\mathbb{R}^n$ with n-dimensional cubes of edge 1 centered in the vertices of the grid. The $|\cdot|_1$-radius of the cells of the tessellation with such cubes is $n/2$. Therefore, claim (i) is a consequence of Theorem 4.1(iii) applied to the space $(\mathbb{R}^n, |\cdot|_1, dx)$ and of Lemma 4.3.

For the proof of (ii) we follow the same argument like in the proof of Theorem 4.1. In this case, we pass from a probability on the grid to a probability on $\mathbb{R}^n$ by averaging on each cube of the tessellation and scaling. Here one should take into account that for a cube C from the tiling

$$\sup\bigl\{|x-y|_1\colon\ x \in C \cap \mathbb{E}^n,\ y \in C\bigr\} = \frac{n+1}{2},$$

that provides the minimal $h = 2(n+1)$ starting from which $h\text{-}\mathbb{C}\mathrm{urv}(\mathbb{E}^n, \mathsf{d}_1, m_n) \geqslant 0$. □

Example 4.4.

(i) Let G be the graph that tiles the euclidian plane with equilateral triangles of edge r. We endow G with the graph metric d_G induced by the euclidian metric and with the 1-dimensional Lebesgue measure m on the edges. Then G has h-curvature $\geqslant 0$ for any $h \geqslant 8r\sqrt{3}/3$.
(ii) The graph G' that tiles the euclidian plane with regular hexagons of edge length r, equipped as usual with the graph metric $\mathsf{d}_{G'}$ and with the 1-dimensional measure m', has h-curvature $\geqslant 0$ for any $h \geqslant 34r/3$.

Proof. Consider a Cartesian coordinate system in the euclidian plane with origin O and axes Ox and Oy. We equip $\mathbb{R}^2$ with the Banach norm $\|\cdot\|$ that has as unit ball the regular hexagon centered in O, having two opposite vertices on Ox and the edge length (measured in the euclidian metric) equal to 1. Explicitly $\|(x,y)\| = \max\{\frac{2\sqrt{3}}{3}|y|, |x| + \frac{\sqrt{3}}{3}|y|\}$ for any (x,y) in $\mathbb{R}^2$. We denote by d the metric determined by this norm.

(i) For the triangular tessellation we choose the origin O to be one of the vertices of the graph and two of the 6 edges emanating from O be along the Ox axis. The edges of the graph have length r in the euclidian metric. We see that $\mathsf{d}_G(v_1, v_2) = \mathsf{d}(v_1, v_2)$ for any two vertices v_1 and v_2 of the graph. In general for $x, y \in G$ we have $|\mathsf{d}_G(x,y) - \mathsf{d}(x,y)| \leqslant r$. Then one can construct a coupling $\hat{\mathsf{d}}$ of d_G and d by setting $\hat{\mathsf{d}}(v,x) := \mathsf{d}(v,x)$ for v vertex of G and $x \in \mathbb{R}^2$ and $\hat{\mathsf{d}}(y,x) := \inf_{i=1,2}\{\mathsf{d}_G(y, v_i) + \mathsf{d}(v_i, x))\}$ if $y \in G$ belongs to an edge with endpoints v_1, v_2 and $x \in \mathbb{R}^2$.

By Lemma 4.3 $\mathbb{C}\mathrm{urv}(\mathbb{R}^2, \mathsf{d}, \lambda) \geqslant 0$, where λ is the 2-dimensional Lebesgue measure. If we tile the plane with regular hexagons A_j, $j \in \mathbb{N}$, which have vertices in the centers of the triangles of the graph G, we have $\hat{\mathsf{d}}(y,x) \leqslant 2r\sqrt{3}/3$ for any $y \in A_j \cap G$ and $x \in A_j$. The proof of the h-curvature bound is a modification of the proof of Theorem 4.1. We start with $\nu_0, \nu_1 \in \mathcal{P}_2^*(G, \mathsf{d}_G, m)$ with $\nu_i = \rho_i m$, $i = 0, 1$, and we define

$$\widetilde{\nu}_i := \sum_{j=1}^{\infty} \frac{1}{\lambda(A_j)} \biggl(\int_{G \cap A_j} \rho_i \, dm \biggr) 1_{A_j} \cdot \lambda \in \mathcal{P}_2^*\bigl(\mathbb{R}^2, \mathsf{d}, \lambda\bigr) \quad \text{for } i = 0, 1.$$

We have then $\hat{\mathsf{d}}_W(\nu_i, \widetilde{\nu_i}) \leqslant 2r\sqrt{3}/3$. We consider $\widetilde{\eta}_t = \widetilde{\rho}_t \cdot \lambda$ the geodesic that joints $\widetilde{\nu}_0$ and $\widetilde{\nu}_1$, along which the convexity condition for the entropy on $\mathcal{P}_2^*(\mathbb{R}^2, \mathsf{d}, \lambda)$ is fulfilled and denote

$$\eta_t := \sum_{j=1}^{\infty} \frac{1}{m(G \cap A_j)} \left(\int_{A_j} \widetilde{\rho}_t \, d\lambda \right) 1_{G \cap A_j} \cdot m.$$

Then η_t is $8r\sqrt{3}/3$-rough t-approximate point between ν_0 and ν_1. From Jensen's inequality we obtain $\mathrm{Ent}(\eta_t \mid m) \leqslant \mathrm{Ent}(\widetilde{\eta}_t \mid \lambda) - \log m(G \cap A) + \log \lambda(A)$ and $\mathrm{Ent}(\widetilde{\nu_i} \mid \lambda) \leqslant \mathrm{Ent}(\nu_i \mid m) + \log m(G \cap A) - \log \lambda(A)$ (observe that all sets A_j have the same Lebesgue measure $\lambda(A)$ and all sets $G \cap A_j$ have the same measure $m(G \cap A)$). Hence η_t satisfies

$$\mathrm{Ent}(\eta_t \mid m) \leqslant (1-t)\,\mathrm{Ent}(\nu_0 \mid m) + t\,\mathrm{Ent}(\nu_1 \mid m),$$

and so we have proved h-$\mathbb{C}\mathrm{urv}(G, \mathsf{d}_G, m) \geqslant 0$ for any $h \geqslant 8r\sqrt{3}/3$.

(ii) For the hexagonal tessellation let O be again one of the vertices of the graph and one of the 3 edges emanating from it be along the Oy axis. In this case we use the Banach norm $\|\cdot\|' := \frac{3}{4}\|\cdot\|$ on $\mathbb{R}^2$ and denote by d' the associated metric. The length of the edges of the graph in the metric d' is equal to $4r/3$. We see that $\mathsf{d}_{G'}(v_1, v_2) = \mathsf{d}'(v_1, v_2)$ for any two vertices v_1, v_2 with $\mathsf{d}_{G'}(v_1, v_2) = 2kr$, $k \in \mathbb{N}$. In general $|\mathsf{d}_{G'} - \mathsf{d}'| \leqslant r/3$ on the set of vertices and $|\mathsf{d}_{G'} - \mathsf{d}'| \leqslant r$ everywhere on G'.

One can construct then a coupling $\hat{\mathsf{d}}'$ of $\mathsf{d}_{G'}+$ and d' in the following way: Fix $v_0 = O$. If v is a vertex of the graph with $\mathsf{d}_{G'}(v_0, v) = 2kr$, $k \in \mathbb{N}$ then set $\hat{\mathsf{d}}'(v, x) := \mathsf{d}'(v, x)$, $x \in \mathbb{R}^2$. For $y \in G'$ with $\mathsf{d}_{G'}(v_0, y) \neq 2kr$, $k \in \mathbb{N}$ define $\hat{\mathsf{d}}'(y, x) := \inf\{\mathsf{d}_{G'}(y, v) + \mathsf{d}'(v, x)\colon\ v \in G',\ \mathsf{d}_{G'}(v_0, v) = 2kr\}$.

We tile the plane with equilateral triangles B_i, $i \in \mathbb{N}$, with vertices in the centers of the hexagons of the graph. Then $\hat{\mathsf{d}}'(y, x) \leqslant 17r/6$ for $y \in B_i \cap G'$, $x \in B_i$. By the same argument as for the triangular tiling we obtain h-$\mathbb{C}\mathrm{urv}(G', \mathsf{d}_{G'}, m') \geqslant 0$ for any $h \geqslant 4 \cdot 17r/6 = 34r/3$. $\square$

5. Some remarks on homogeneous planar graphs

We refer in the sequel to a special class of graphs. In general, a graph G is determined by the set of vertices $V(G)$ and the set of edges $E(G)$. In order to regard graphs as discrete analogues of 2-dimensional manifolds one has to specify also the set of faces $F(G)$ and to impose the graph to be planar. A graph is planar if it can be drawn in a plane without graph edges crossing (i.e., it has graph crossing number 0). Only planar graphs have duals. The graphs we will be concerned with are connected and simple (with no self-loops and no multiple edges) and such that their dual graphs are also simple, therefore any two faces have at most one common edge and every face is bounded by a cycle.

We consider in the following the (possibly infinite) *homogeneous graph* $\mathbb{G}(l, n, r)$ with vertices of constant degree $l \geqslant 3$, with faces bounded by polygons with $n \geqslant 3$ edges (thus n is the degree of all vertices in the dual graph) and such that all edges have the same length $r > 0$ (see Fig. 1).

The following result is probably well-known, but since we did not find a reference we present here the easy proof.

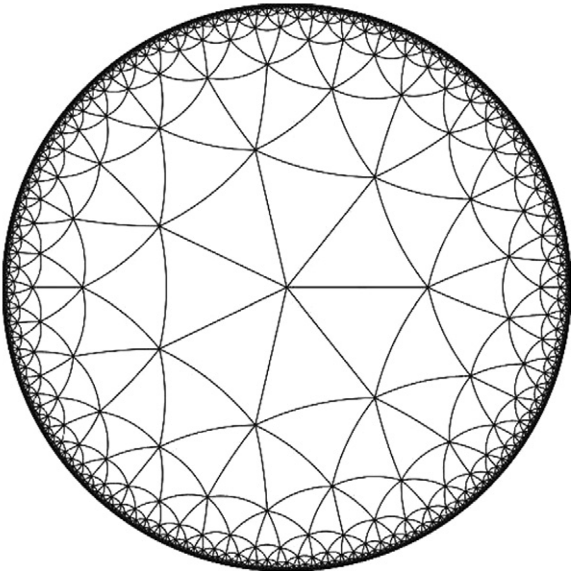

Fig. 1. $\mathbb{G}(7,3,r)$.

Lemma 5.1.

(i) *If $\frac{1}{l}+\frac{1}{n}<\frac{1}{2}$ then $\mathbb{G}(l,n,r)$ can be embedded into the 2-dimensional hyperbolic space with constant sectional curvature*

$$K=-\frac{1}{r^2}\left[\operatorname{arccosh}\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})}-1\right)\right]^2. \tag{5.1}$$

There are infinitely many choices of such l and n. In any case, the graph is unbounded.

(ii) *If $\frac{1}{l}+\frac{1}{n}>\frac{1}{2}$ then $\mathbb{G}(l,n,r)$ is one of the five regular polyhedra (Tetrahedron, Octahedron, Cube, Icosahedron, Dodecahedron) and can be embedded into the 2-dimensional sphere with constant sectional curvature*

$$K=\frac{1}{r^2}\left[\arccos\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})}-1\right)\right]^2. \tag{5.2}$$

(iii) *If $\frac{1}{l}+\frac{1}{n}=\frac{1}{2}$ then $\mathbb{G}(l,n,r)$ can be embedded into the euclidian plane ($K=0$). In this case there are exactly three cases corresponding to the 3 regular tessellations of the euclidian plane: the tessellation of triangles ($l=6$, $n=3$), of squares ($l=n=4$), and of hexagons ($l=3$, $n=6$).*

Proof. Firstly we see that

$$2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})}-1>1 \quad\Leftrightarrow\quad \sin^2\left(\frac{\pi}{2}-\frac{\pi}{n}\right)>\sin^2\left(\frac{\pi}{l}\right) \quad\Leftrightarrow\quad \frac{1}{l}+\frac{1}{n}<\frac{1}{2}$$

hence in each case the expression that defines the curvature K makes sense.

(i) For given l, n, r we construct the embedding in the following way: we start from an arbitrary point O of the 2-hyperbolic space with curvature K, denoted by $\mathbb{H}^{K,2}$. From this point we construct n geodesic lines $OA_1, OA_2, \dots, OA_n$ of length

$$R := \frac{1}{\sqrt{-K}} \operatorname{arcsinh}\left(\frac{\sinh(\sqrt{-K}r)}{\sin(\frac{2\pi}{n})} \sin\left(\frac{\pi}{l}\right)\right), \tag{5.3}$$

such that the inner angle between any two consecutive geodesics OA_k, OA_{k+1} is $2\pi/n$. We prove that $A_1, A_2, \dots, A_n$ correspond to vertices of the given graph, and the geodesics $A_1A_2, \dots, A_{n-1}A_n, A_nA_1$ correspond isometrically to consecutive edges in $\mathbb{G}(l,n,r)$ that bound a regular n-polygon with edge-length r and all angles equal to $2\pi/l$. Let us denote by d the intrinsic metric on $\mathbb{H}^{K,2}$.

From the Cosine Rule for hyperbolic triangles applied to ΔOA_1A_2 and from (5.1) and (5.3) we have:

$$\begin{aligned}
\cosh\big(\sqrt{-K}\mathsf{d}(A_1,A_2)\big) &= \cosh^2(\sqrt{-K}R) - \sinh^2(\sqrt{-K}R)\cos\left(\frac{2\pi}{n}\right) \\
&= 1 + \sinh^2(\sqrt{-K}R)\left(1 - \cos\left(\frac{2\pi}{n}\right)\right) \\
&= 1 + \frac{\sinh^2(\sqrt{-K}r)}{\sin^2(\frac{2\pi}{n})} \sin^2\left(\frac{\pi}{l}\right)\left(1 - \cos\left(\frac{2\pi}{n}\right)\right) \\
&= 1 + \frac{\cosh^2(\sqrt{-K}r) - 1}{1 + \cos(\frac{2\pi}{n})} \sin^2\left(\frac{\pi}{l}\right) \\
&= 1 + \frac{\sin^2(\frac{\pi}{l})}{2\cos^2(\frac{\pi}{n})}\left[\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1\right)^2 - 1\right] \\
&= 2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1 = \cosh(\sqrt{-K}r),
\end{aligned}$$

so $\mathsf{d}(A_1, A_2) = r$ and the same holds for all the other edges of the polygon. We apply now the Sine Rule for the hyperbolic triangle ΔOA_1A_2 and (5.3) in order to compute:

$$\sin \sphericalangle(A_1; O, A_2) = \frac{\sin(\frac{2\pi}{n})}{\sinh(\sqrt{-K}r)} \sinh(\sqrt{-K}R) = \sin\left(\frac{\pi}{l}\right), \tag{5.4}$$

where $\sphericalangle(A_1; O, A_2)$ denotes the angle at A_1 in the triangle ΔOA_1A_2. This angle is less then $\pi/2$ because it is equal to $\sphericalangle(A_2; O, A_1)$ and in the hyperbolic triangles the sum of the angles of a triangle is less then π. Therefore (5.4) shows that all the angles of the polygon are equal to $2\pi/l$, so around each vertex one can construct other $l-1$ polygons with n edges, congruent with the first one. We repeat the procedure with each of the vertices of the new polygons. In this way the whole space $\mathbb{H}^{K,2}$ can be tiled with regular polygons which are faces of the graph $\mathbb{G}(l,n,r)$.

(ii), (iii) Since there is only a finite number of examples with well-known realizations, the claim can be verified directly. Alternatively, one can prove it like in the part (i) with appropriate interpretations of the hyperbolic sine as sine for positive curvature and as length for the euclidian plane. □

Remark 5.2. The dual graph $\mathbb{G}(l,n,r)^* = \mathbb{G}(n,l,r')$ is embedded into the 2-manifold of the same constant curvature as $\mathbb{G}(l,n,r)$, where the dual edge length is

$$r' := r \cdot \operatorname{arccosh}\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1\right) \Big/ \operatorname{arccosh}\left(2\frac{\cos^2(\frac{\pi}{l})}{\sin^2(\frac{\pi}{n})} - 1\right) \quad \text{for } K < 0,$$

and with appropriate modifications for the other two cases.

In each of the three cases from Lemma 5.1 the 2-manifold will be endowed with the intrinsic metric d and with the Riemannian volume vol. We equip $\mathbb{G}(l,n,r)$ with the metric d induced by the corresponding Riemannian metric and with the uniform measure m on the edges. We denote further by $\mathbb{V}(l,n,r)$ the set of vertices of the graph $\mathbb{G}(l,n,r)$ equipped with the same metric d inherited from the Riemannian manifold and with the counting measure $\widetilde{m} := \sum_{v\in\mathbb{V}} \delta_v$.

Theorem 5.3. *For any numbers $l,n \geqslant 3$ and for any $r > 0$ both metric measure spaces $(\mathbb{V}(l,n,r), \mathsf{d}, \widetilde{m})$ and $(\mathbb{G}(l,n,r), \mathsf{d}, m)$ have h-curvature $\geqslant K$ for $h \geqslant r \cdot C(l,n)$, where*

$$K = \begin{cases} -\frac{1}{r^2}\left[\operatorname{arccosh}\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1\right)\right]^2 & \text{for } \frac{1}{l} + \frac{1}{n} > \frac{1}{2}, \\ \frac{1}{r^2}\left[\arccos\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1\right)\right]^2 & \text{for } \frac{1}{l} + \frac{1}{n} < \frac{1}{2}, \\ 0 & \text{for } \frac{1}{l} + \frac{1}{n} = \frac{1}{2} \end{cases} \tag{5.5}$$

and $C(l,n) = 4 \cdot \operatorname{arcsinh}\left(\frac{1}{sin(\frac{\pi}{n})}\sqrt{\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1}\right) / \operatorname{arccosh}\left(2\frac{\cos^2(\frac{\pi}{n})}{\sin^2(\frac{\pi}{l})} - 1\right)$.

Proof. We look at $\mathbb{V}(l,n,r)$ and $\mathbb{G}(l,n,r)$ as subsets of the 2-manifold with constant curvature K (given by Lemma 5.1). We tile the manifold with the faces of the dual graph $\mathbb{G}(n,l,r')$ having vertices in the centers of the faces of $\mathbb{G}(l,n,r)$ (the center O of the polygon with n edges in the proof of Lemma 5.1 becomes vertex of the dual).

We make explicitly the calculations only in the hyperbolic case, the other two cases are similar. One can decompose the hyperbolic space as $\mathbb{H}^{K,2} = \bigcup_{j=1}^{\infty} F_j$, where $\{F_j\}_j$ are the faces of the dual graph, as described above. The curvature bound for the discrete space $\mathbb{V}(l,n,r)$ is then a consequence of the Theorem 4.1. For $\mathbb{G} := \mathbb{G}(l,n,r)$ the proof of the curvature bound is a modification of the proof of Theorem 4.1. We start with $\nu_0, \nu_1 \in \mathcal{P}_2^*(\mathbb{G}(l,n,r), \mathsf{d}, m)$ with $\nu_i = \rho_i \cdot m$, $i = 0, 1$, and define

$$\widetilde{\nu_i} := \sum_{j=1}^{\infty} \frac{1}{\operatorname{vol}(F_j)} \left(\int_{\mathbb{G}\cap F_j} \rho_i \, dm \right) 1_{F_j} \cdot \operatorname{vol} \in \mathcal{P}_2^*\left(\mathbb{H}^{K,2}, \mathsf{d}, \operatorname{vol}\right) \quad \text{for } i = 0, 1.$$

Now the place of $R(h)$ from Theorem 4.1 is taken by R from the proof of Lemma 5.1(i), so $\mathsf{d}_W(\nu_i, \widetilde{\nu_i}) \leqslant R$. One can express R only in terms of our initial data l, n and r as $R = rC(l,n)/4$, with $C(l,n)$ given in the statement of the theorem. We consider $\widetilde{\eta_t} = \widetilde{\rho_t} \cdot \operatorname{vol}$ the geodesic that joints $\widetilde{\nu_0}$ and $\widetilde{\nu_1}$, along which one has the K-convexity for the entropy on $\mathbb{H}^{K,2}$ (Theorem 4.9 from [12]) and denote

$$\eta_t := \sum_{j=1}^{\infty} \frac{1}{m(\mathbb{G} \cap F_j)} \left(\int_{F_j} \widetilde{\rho}_t \, d\,\mathrm{vol} \right) 1_{\mathbb{G} \cap F_j} \cdot m.$$

Then η_t is $4R$-rough t-approximate point between ν_0 and ν_1. From Jensen's inequality we obtain $\mathrm{Ent}(\eta_t \mid m) \leqslant \mathrm{Ent}(\widetilde{\eta_t} \mid \mathrm{vol}) - \log m(\mathbb{G} \cap F) + \log \mathrm{vol}(F)$ and $\mathrm{Ent}(\widetilde{\nu_i} \mid \mathrm{vol}) \leqslant \mathrm{Ent}(\nu_i \mid m) + \log m(\mathbb{G} \cap F) - \log \mathrm{vol}(F)$ (observe that all faces F_j have the same volume $\mathrm{vol}(F)$ and all sets $\mathbb{G} \cap F_j$ have the same measure $m(\mathbb{G} \cap F)$). Hence, like in the proof of Theorem 4.1, η_t satisfies

$$\mathrm{Ent}(\eta_t \mid m) \leqslant (1-t)\,\mathrm{Ent}(\nu_0 \mid m) + t\,\mathrm{Ent}(\nu_1 \mid m) - \frac{K}{2} t(1-t) \mathsf{d}_W^{-2R}(\nu_0, \nu_1)^2,$$

so we have proved h-$\mathbb{C}\mathrm{urv}(\mathbb{G}(l,n,r), \mathsf{d}, m) \geqslant K$ for any $h \geqslant 4R$ in the hyperbolic case $(K < 0)$. □

Remark 5.4. There are various notions of combinatorial curvature for graphs in the literature, see for instance [3,5,6]. The notion of curvature introduced by Gromov in [5] was used in studying hyperbolic groups. Later on it was modified and investigated by Higuchi [6] and other authors. Forman has introduced in [3] a different notion of combinatorial Ricci curvature for cell complexes. The graphs considered in the above mentioned works have neither specified metric, nor specified reference measure.

In [6] the combinatorial curvature of a graph G is a map $\Phi_G : V(G) \to \mathbb{R}$ that assigns to each vertex $x \in V(G)$ the number $\Phi_G(x) = 1 - \frac{m(x)}{2} + \sum_{i=1}^{m(x)} \frac{1}{d(F_i)}$, where $m(x)$ is the degree of the vertex x, $d(F)$ is the number of edges of the cycle bounding a face F, and $F_1, F_2, \ldots, F_{m(x)}$ are the faces around the vertex x. The combinatorial curvature introduced in [5] is a map $\Phi_G^* : F(G) \to \mathbb{R}$, where the curvature $\Phi_G^*(F)$ of a face F is given by the curvature Φ_G of the corresponding vertex in the dual graph. For the homogeneous graph $\mathbb{G}(l,n,r)$, the curvature of any vertex x is $\Phi_G(x) = l(\frac{1}{l} + \frac{1}{n} - \frac{1}{2})$ and the curvature in the sense of Gromov [5] of any face F is $\Phi_G^*(F) = n(\frac{1}{l} + \frac{1}{n} - \frac{1}{2})$.

Note that the sign of the combinatorial curvature in both approaches above changes according to whether $\frac{1}{l} + \frac{1}{n}$ is greater or less than $\frac{1}{2}$. Rather curiously, in our Theorem 5.3 the sign of the rough curvature bound changes in the same manner, although our notion of curvature applies to graphs that have a metric structure and a reference measure. For the moment we see no further links with the notions of combinatorial curvature mentioned here.

6. Perturbed transportation inequalities, concentration of measure and exponential integrability

Let (M,d) be a metric space and $m \in \mathcal{P}_2(M, \mathsf{d})$ be a given probability measure. The measure m is said to satisfy a Talagrand inequality (or a transportation cost inequality) with constant K iff for all $\nu \in \mathcal{P}_2(M, \mathsf{d})$

$$\mathsf{d}_W(\nu, m) \leqslant \sqrt{\frac{2\,\mathrm{Ent}(\nu \mid m)}{K}}. \tag{6.1}$$

Such an inequality was first proved by Talagrand in [13] for the canonical Gaussian measure on $\mathbb{R}^n$. A positive rough curvature bound allows us to obtain a weaker inequality, in terms of the perturbation d_W^{+h} of the Wasserstein distance:

Proposition 6.1 *("h-Talagrand inequality"). Assume that (M, d, m) is a metric measure space which has h-$\mathbb{C}$urv$(M, \mathsf{d}, m) \geqslant K$ for some numbers $h > 0$ and $K > 0$. Then for each $\nu \in \mathcal{P}_2(M, \mathsf{d})$ we have*

$$\mathsf{d}_W^{+h}(\nu, m) \leqslant \sqrt{\frac{2\,\mathrm{Ent}(\nu \mid m)}{K}}. \tag{6.2}$$

We will call (6.2) *h-Talagrand inequality.*

Proof. Since we assumed that m is a probability measure, for any $\nu \in \mathcal{P}_2(M, \mathsf{d})$ the entropy functional is nonnegative: $\mathrm{Ent}(\nu \mid m) \geqslant -\log m(M) = 0$, according to Lemma 4.1 from [12]. The curvature bound h-$\mathbb{C}\mathrm{urv}(M, \mathsf{d}, m) \geqslant K$ implies that for the pair of measures ν and m and for each $t \in [0, 1]$ there exists an h-rough t-approximate point $\eta_t \in \mathcal{P}_2(M, \mathsf{d})$ such that

$$\mathrm{Ent}(\eta_t \mid m) \leqslant (1-t)\,\mathrm{Ent}(\nu \mid m) - \frac{K}{2} t(1-t)\mathsf{d}_W^{+h}(\nu, m)^2. \tag{6.3}$$

If $\mathrm{Ent}(\nu \mid m) < \frac{K}{2}\mathsf{d}_W^{+h}(\nu, m)^2$ then there exists an $\epsilon > 0$ such that $\mathrm{Ent}(\nu \mid m) + \epsilon < \frac{K}{2}\mathsf{d}_W^{+h}(\nu, m)^2$. This together with (6.3) would imply

$$\mathrm{Ent}(\eta_t \mid m) < \frac{K}{2}(1-t)^2\mathsf{d}_W^{+h}(\nu, m)^2 - \epsilon(1-t)$$

for each $t \in [0, 1]$. We choose now t very close to 1, such that $0 < 1 - t < \epsilon$ and $K(1-t)^2\mathsf{d}_W^{+h}(\nu, m)^2 < \epsilon^2$. This entails $\mathrm{Ent}(\eta_t \mid m) < -\epsilon^2/2 < 0$, in contradiction with the fact that the entropy functional is nonnegative. Therefore $\mathrm{Ent}(\nu \mid m) \geqslant \frac{K}{2}\mathsf{d}_W^{+h}(\nu, m)^2$, which is precisely our claim. □

A Talagrand inequality for the measure m implies a concentration of measure inequality for m (see for instance [9]).

For a given Borel set $A \subset M$ denote the (open) r-neighborhood of A by $B_r(A) := \{x \in M\colon\ \mathsf{d}(x, A) < r\}$ for $r > 0$. The concentration function of (M, d, m) is defined as

$$\alpha_{(M,\mathsf{d},m)}(r) := \sup\left\{1 - m\big(B_r(A)\big)\colon\ A \in \mathcal{B}(M),\ m(A) \geqslant \frac{1}{2}\right\}, \quad r > 0.$$

We refer to [7] for further details on measure concentration.

The following result shows that positive rough curvature bound implies a normal concentration inequality, via h-Talagrand inequality.

Proposition 6.2. *Let (M, d, m) be a metric measure space with h-$\mathbb{C}$urv$(M, \mathsf{d}, m) \geqslant K > 0$ for some $h > 0$. Then there exists an $r_0 > 0$ such that for all $r \geqslant r_0$*

$$\alpha_{(M,\mathsf{d},m)}(r) \leqslant e^{-Kr^2/8}.$$

Proof. We follow essentially the argument of K. Marton used in [9] for obtaining concentration of measure out of a Talagrand inequality for the Wasserstein distance of order 1. Let $A, B \in \mathcal{B}(M)$ be given with $m(A), m(B) > 0$. Consider the conditional probabilities $m_A = m(\cdot \mid A)$ and $m_B = m(\cdot \mid B)$. For these measures the h-Talagrand inequality holds:

$$\mathsf{d}_W^{+h}(m_A, m) \leqslant \sqrt{\frac{2\,\mathrm{Ent}(m_A \mid m)}{K}}, \qquad \mathsf{d}_W^{+h}(m_B, m) \leqslant \sqrt{\frac{2\,\mathrm{Ent}(m_B \mid m)}{K}}. \tag{6.4}$$

Let q_A and q_B be the $+h$-optimal couplings of m_A, m and m_B, m respectively. According to [2], section 11.8, there exists a probability measure $\hat{q}$ on $M \times M \times M$ such that its projection on the first two factors is q_A and the projection on the last two factors is q_B. Then we have in turn

$$\begin{aligned}
&\mathsf{d}_W^{+h}(m_A, m) + \mathsf{d}_W^{+h}(m, m_B) \\
&\quad = \left\{ \int\limits_{M\times M\times M} \left[\left(\mathsf{d}(x_1, x_2) - h\right)_+\right]^2 d\hat{q}(x_1, x_2, x_2) \right\}^{1/2} \\
&\qquad + \left\{ \int\limits_{M\times M\times M} \left[\left(\mathsf{d}(x_2, x_3) - h\right)_+\right]^2 d\hat{q}(x_1, x_2, x_2) \right\}^{1/2} \\
&\quad \geqslant \left\{ \int\limits_{M\times M\times M} \left[\left(\mathsf{d}(x_1, x_2) - h\right)_+ + (\mathsf{d}(x_2, x_3) - h)_+\right]^2 d\hat{q}(x_1, x_2, x_2) \right\}^{1/2} \\
&\quad \geqslant \left\{ \int\limits_{M\times M\times M} \left[\left(\mathsf{d}(x_1, x_2) + \mathsf{d}(x_2, x_3) - 2h\right)_+\right]^2 d\hat{q}(x_1, x_2, x_2) \right\}^{1/2} \\
&\quad \geqslant \left\{ \int\limits_{M\times M\times M} \left[\left(\mathsf{d}(x_1, x_3) - 2h\right)_+\right]^2 d\hat{q}(x_1, x_2, x_2) \right\}^{1/2}.
\end{aligned}$$

Assume now that $\mathsf{d}(A, B) \geqslant 2h$. Since the projection on the first factor of $\hat{q}$ is m_A and the projection on the last factor is m_B, the support of $\hat{q}$ must be a subset of $A \times M \times B$, hence

$$\left\{ \int\limits_{M\times M\times M} \left[\left(\mathsf{d}(x_1, x_3) - 2h\right)_+\right]^2 d\hat{q}(x_1, x_2, x_2) \right\}^{1/2} \geqslant \mathsf{d}(A, B) - 2h.$$

The above estimates together with (6.4) imply

$$\begin{aligned}
\mathsf{d}(A, B) - 2h &\leqslant \sqrt{\frac{2\ \mathrm{Ent}(m_A \mid m)}{K}} + \sqrt{\frac{2\ \mathrm{Ent}(m_B \mid m)}{K}} \\
&= \sqrt{\frac{2}{K} \log \frac{1}{m(A)}} + \sqrt{\frac{2}{K} \log \frac{1}{m(B)}}.
\end{aligned}$$

If we choose now $2h \leqslant r$ and for a given $A \in \mathcal{B}(M)$ we replace B by $\complement B_r(A)$, we get

$$r - 2h \leqslant \sqrt{\frac{2}{K} \log \frac{1}{m(A)}} + \sqrt{\frac{2}{K} \log \frac{1}{1 - m(B_r(A))}}.$$

Hence, for $m(A) \geqslant \frac{1}{2}$

$$r - 2h \leqslant \sqrt{\frac{2}{K} \log 2} + \sqrt{\frac{2}{K} \log \frac{1}{1 - m(B_r(A))}}.$$

Therefore whenever $r \geqslant 2\sqrt{\frac{2}{K} \log 2} + 4h$ for instance we have

$$\frac{r}{2} \leqslant \sqrt{\frac{2}{K} \log \frac{1}{1 - m(B_r(A))}},$$

or equivalently

$$1 - m\big(B_r(A)\big) \leqslant e^{-Kr^2/8},$$

which ends the proof. □

In [1] it has been shown that a Talagrand type inequality implies exponential integrability of the Lipshitz functions. We prove further that an h-Talagrand inequality leads to the same conclusion.

Theorem 6.3. *Assume that (M, d) is a metric space and let $h > 0$ be given. If m is a probability measure on (M, d) that satisfies an h-Talagrand inequality of constant $K > 0$ then all Lipschitz functions are exponentially integrable. More precisely, for any Lipschitz function φ with $\|\varphi\|_{\mathrm{Lip}} \leqslant 1$ and $\int \varphi \, dm = 0$ we have*

$$\forall t > 0, \quad \int_M e^{t\varphi} \, dm \leqslant e^{\frac{t^2}{2K} + ht}, \tag{6.5}$$

or equivalently, for any Lipschitz function φ

$$\forall t > 0, \quad \int_M e^{t\varphi} \, dm \leqslant \exp\left(t \int_M \varphi \, dm\right) \exp\left(\frac{t^2}{2K} \|\varphi\|_{\mathrm{Lip}}^2 + ht \|\varphi\|_{\mathrm{Lip}}\right). \tag{6.6}$$

Proof. The proof we present here extends the one given in [1]. Let f be a probability density with $f \log f$ integrable with respect to m. The h-Talagrand inequality implies

$$\mathsf{d}_W^{+h}(fm, m) \leqslant \sqrt{\frac{2}{K} \int_M f \log f \, dm} \leqslant \frac{t}{2K} + \frac{1}{t} \int_M f \log f \, dm$$

for each $t > 0$. We consider now the Wasserstein distance of order 1 of two probability measures μ and ν

$$\mathsf{d}_W^1(\mu, \nu) := \inf \int\limits_{M\times M} \mathsf{d}(x_0, x_1)\, dq(x_0, x_1),$$

where q ranges over all couplings of μ and ν. If $\widetilde{q}$ is a $+h$-optimal coupling of fm and m then by the Cauchy–Schwartz inequality,

$$\begin{aligned}\mathsf{d}_W^{+h}(fm, m) &= \left\{ \int\limits_{M\times M} \left[\left(\mathsf{d}(x_0, x_1) - h\right)_+\right]^2 d\widetilde{q}(x_0, x_1)\right\}^{1/2} \\ &\geqslant \int\limits_{M\times M} \left(\mathsf{d}(x_0, x_1) - h\right)_+ d\widetilde{q}(x_0, x_1) \geqslant \mathsf{d}_W^1(fm, m) - h.\end{aligned}$$

The Kantorovich–Rubinstein theorem gives the following duality formula

$$\mathsf{d}_W^1(fm, m) = \sup_{\|\varphi\|_{\mathrm{Lip}}\leqslant 1} \left\{ \int\limits_M \varphi f\, dm - \int\limits_M \varphi\, dm \right\}.$$

If φ is a Lipschitz function that satisfies the assumptions of the theorem ($\|\varphi\|_{\mathrm{Lip}} \leqslant 1$ and $\int \varphi\, dm = 0$) then

$$\int\limits_M \varphi f\, dm \leqslant \mathsf{d}_W^{+h}(fm, m) + h \leqslant \frac{t}{2K} + \frac{1}{t} \int\limits_M f \log f\, dm + h,$$

which can be written as

$$\int\limits_M \left(t\varphi - \frac{t^2}{2K}\right) f\, dm \leqslant \int\limits_M f \log f\, dm + ht. \tag{6.7}$$

This estimate should take place for any probability density f. Therefore one can take

$$f = e^{t\varphi - \frac{t^2}{2K}} \left(\int\limits_M e^{t\varphi - \frac{t^2}{2K}}\, dm \right)^{-1}$$

in formula (6.7) and obtain

$$\begin{aligned}&\left\{ \int\limits_M \left(t\varphi - \frac{t^2}{2K}\right) e^{t\varphi - \frac{t^2}{2K}}\, dm \right\} \left(\int\limits_M e^{t\varphi - \frac{t^2}{2K}}\, dm \right)^{-1} \\ &\quad\leqslant \int\limits_M e^{t\varphi - \frac{t^2}{2K}} \left(\int\limits_M e^{t\varphi - \frac{t^2}{2K}}\, dm \right)^{-1} \left\{ t\varphi - \frac{t^2}{2K} - \log\left(\int\limits_M e^{t\varphi - \frac{t^2}{2K}}\, dm \right) \right\} dm + ht.\end{aligned}$$

This yields

$$\log\left(\int_M e^{t\varphi-\frac{t^2}{2K}}\,dm\right)dm \leqslant ht,$$

that proves the claim (6.5). The general estimate (6.6) is a consequence of (6.5) applied to the function $\psi = \frac{1}{\|\varphi\|_{\mathrm{Lip}}}[\varphi - \int \varphi\,dm]$. □

Remark 6.4. In the continuous case, by formal calculus, the following two assertions are equivalent (see [11] for the case of Riemannian manifolds):

(i) The entropy functional $\mathrm{Ent}(\cdot \mid m)$ is weakly K-convex on $\mathcal{P}_2(M, \mathsf{d})$, in the sense of inequality (2.1);
(ii) The gradient flow $\Phi : \mathbb{R}_+ \times \mathcal{P}_2(M, \mathsf{d}) \to \mathcal{P}_2(M, \mathsf{d})$ with respect to $\mathrm{Ent}(\cdot \mid m)$ satisfies

$$\mathsf{d}_W\big(\Phi(t,\mu), \Phi(t,\nu)\big) \leqslant e^{-Kt}\mathsf{d}_W(\mu,\nu) \quad \forall \mu,\nu \in \mathcal{P}_2(M,\mathsf{d}),\ \forall t \geqslant 0. \tag{6.8}$$

The rough notion of curvature bound that we have introduced in this paper is a discrete version of (2.1), whereas the approach presented in [10] is a discrete form of (6.8). Both imply e.g. measure concentration, although in general there is no real overlap, since in the discrete case there is no direct relation between Markov chains and entropy functionals.

References

[1] S. Bobkov, F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, J. Funct. Anal. 163 (1999) 1–28.
[2] R.M. Dudley, Real Analysis and Probability, The Wadsworth Brooks/Cole Mathematics Series, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1989, reprinted by Cambridge University Press, 2002.
[3] R. Forman, Bochner's method for cell complexes and combinatorial Ricci curvature, Discrete Comput. Geom. 29 (2003) 323–374.
[4] K. Fukaya, Collapsing of Riemannian manifolds and eigenvalues of Laplace operator, Invent. Math. 87 (1987) 517–547.
[5] M. Gromov, Hyperbolic groups, in: Essays in Group Theory, in: Math. Sci. Res. Inst. Publ., vol. 8, Springer, 1987, pp. 75–263, New York.
[6] Y. Higuchi, Combinatorial curvature for planar graphs, J. Graph Theory 38 (2001) 220–229.
[7] M. Ledoux, The Concentration of Measure Phenomenon, Math. Surveys Monogr., vol. 89, Amer. Math. Soc., 2001.
[8] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. of Math. 169 (2009), in press.
[9] K. Marton, A measure concentration inequality for contracting Markov chains, Geom. Funct. Anal. 6 (1997) 556–571.
[10] Y. Ollivier, Ricci curvature of Markov chains on metric spaces, J. Funct. Anal. 256 (2009) 810–864.
[11] M.K. von Renesse, K.T. Sturm, Transport inequalities, gradient estimates, entropy, and Ricci curvature, Comm. Pure Appl. Math. 58 (2005) 923–940.
[12] K.T. Sturm, On the geometry of metric measure spaces. I, Acta Math. 196 (2006) 65–131.
[13] M. Talagrand, Transportation cost for Gaussian and other product measures, Geom. Funct. Anal. 6 (1996) 587–600.
[14] C. Villani, Topics in Mass Transportation, Grad. Stud. Math., Amer. Math. Soc., 2003.
[15] C. Villani, Optimal Transport: Old and New, Grundlehren Math. Wiss., vol. 338, Springer, 2008.