SNOWFLAKE UNIVERSALITY OF WASSERSTEIN SPACES

ALEXANDR ANDONI, ASSAF NAOR, AND OFER NEIMAN

Abstract. For $p \in (1,\infty)$ let $\mathcal{P}_p(\mathbb{R}^3)$ denote the metric space of all p-integrable Borel probability measures on $\mathbb{R}^3$, equipped with the Wasserstein p metric W_p. We prove that for every $\varepsilon > 0$, every $\theta \in (0,1/p]$ and every finite metric space (X,d_X), the metric space (X,d_X^θ) embeds into $\mathcal{P}_p(\mathbb{R}^3)$ with distortion at most $1+\varepsilon$. We show that this is sharp when $p \in (1,2]$ in the sense that the exponent $1/p$ cannot be replaced by any larger number. In fact, for arbitrarily large $n \in \mathbb{N}$ there exists an n-point metric space (X_n,d_n) such that for every $\alpha \in (1/p,1]$ any embedding of the metric space (X_n,d_n^α) into $\mathcal{P}_p(\mathbb{R}^3)$ incurs distortion that is at least a constant multiple of $(\log n)^{\alpha-1/p}$. These statements establish that there exists an Alexandrov space of nonnegative curvature, namely $\mathcal{P}_2(\mathbb{R}^3)$, with respect to which there does not exist a sequence of bounded degree expander graphs. It also follows that $\mathcal{P}_2(\mathbb{R}^3)$ does not admit a uniform, coarse, or quasisymmetric embedding into any Banach space of nontrivial type. Links to several longstanding open questions in metric geometry are discussed, including the characterization of subsets of Alexandrov spaces, existence of expanders, the universality problem for $\mathcal{P}_2(\mathbb{R}^k)$, and the metric cotype dichotomy problem.

1. Introduction

We shall start by quickly recalling basic notation and terminology from the theory of transportation cost metrics; all the necessary background can be found in [94]. For a complete separable metric space (X,d_X) and $p \in (0,\infty)$, let $\mathcal{P}_p(X)$ denote the space of all Borel probability measures μ on X satisfying

$$\int_X d_X(x,x_0)^p \mathrm{d}\mu(x) < \infty$$

for some (hence all) $x_0 \in X$. A coupling of a pair of Borel probability measures (μ,ν) on X is a Borel probability measure π on $X \times X$ such that $\mu(A) = \pi(A \times X)$ and $\nu(A) = \pi(X \times A)$ for every Borel measurable $A \subseteq X$. The set of couplings of (μ,ν) is denoted $\Pi(\mu,\nu)$. The Wasserstein p distance between $\mu,\nu \in \mathcal{P}_p(X)$ is defined to be

$$\mathsf{W}_p(\mu,\nu) \stackrel{\mathrm{def}}{=} \inf_{\pi \in \Pi(\mu,\nu)} \left(\iint_{X\times X} d_X(x,y)^p \mathrm{d}\pi(x,y) \right)^{\frac{1}{p}}.$$

W_p is a metric on $\mathcal{P}_p(x)$ whenever $p \geqslant 1$. The metric space $(\mathcal{P}_p(X),\mathsf{W}_p)$ is called the Wasserstein p space over (X,d_X). Unless stated otherwise, in the ensuing discussion whenever we refer to the metric space $\mathcal{P}_p(X)$ it will be understood that $\mathcal{P}_p(X)$ is equipped with the metric W_p.

1.1. **Bi-Lipschitz Embeddings.** Suppose that (X,d_X) and (Y,d_Y) are metric spaces and that $D \in [1,\infty]$. A mapping $f : X \to Y$ is said to have distortion at most D if there exists $s \in (0,\infty)$ such that every $x,y \in X$ satisfy $sd_X(x,y) \leqslant d_Y(f(x),f(y)) \leqslant Dsd_X(x,y)$. The infimum over those $D \in [1,\infty]$ for which this holds true is called the distortion of f and is denoted $\mathbf{dist}(f)$. If there exists a mapping $f : X \to Y$ with distortion at most D then we say that (X,d_X) embeds with

2010 *Mathematics Subject Classification.* 46B85, 53C23, 46E27.
Key words and phrases. Metric embeddings, Wasserstein spaces, Alexandrov spaces, Snowflakes of metric spaces, nonlinear spectral gaps, metric cotype, Markov type.

distortion D into (Y, d_Y). The infimum of $\mathbf{dist}(f)$ over all $f : X \to Y$ is denoted $c_{(Y,d_Y)}(X, d_X)$, or $c_Y(X)$ if the metrics are clear from the context.

1.2. **Snowflake universality.** Below, unless stated otherwise, $\mathbb{R}^n$ will be endowed with the standard Euclidean metric. Here we show that $\mathcal{P}_p(\mathbb{R}^3)$ exhibits the following universality phenomenon.

Theorem 1. *If $p \in (1, \infty)$ then for every finite metric space (X, d_X) we have*

$$c_{(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)}\Big(X, d_X^{\frac{1}{p}}\Big) = 1.$$

For a metric space (X, d_X) and $\theta \in (0, 1]$, the metric space (X, d_X^θ) is commonly called the θ-*snowflake* of (X, d_X); see e.g. [21]. Thus Theorem 1 asserts that the θ-snowflake of any finite metric space (X, d_X) embeds with distortion $1+\varepsilon$ into $\mathcal{P}_p(\mathbb{R}^3)$ for every $\varepsilon \in (0,\infty)$ and $\theta \in (0, 1/p]$ (formally, Theorem 1 makes this assertion when $\theta = 1/p$, but for general $\theta \in (0, 1/p]$ one can then apply Theorem 1 to the metric space $(X, d_X^{\theta p})$ to deduce the seemingly more general statement).

Theorem 2 below implies that Theorem 1 is sharp if $p \in (1, 2]$, and yields a nontrivial, though probably non-sharp, restriction on the embeddability of snowflakes into $\mathcal{P}_p(\mathbb{R}^3)$ also for $p \in (2,\infty)$.

Theorem 2. *For arbitrarily large $n \in \mathbb{N}$ there exists an n-point metric space (X_n, d_{X_n}) such that for every $\alpha \in (0, 1]$ we have*

$$c_{(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)}(X_n, d_{X_n}^\alpha) \gtrsim \begin{cases} (\log n)^{\alpha-\frac{1}{p}} & \text{if } p \in (1, 2], \\ (\log n)^{\alpha+\frac{1}{p}-1} & \text{if } p \in (2,\infty). \end{cases}$$

Here, and in what follows, we use standard asymptotic notation, i.e., for $a, b \in [0,\infty)$ the notation $a \gtrsim b$ (respectively $a \lesssim b$) stands for $a \geqslant cb$ (respectively $a \leqslant cb$) for some universal constant $c \in (0,\infty)$. The notation $a \asymp b$ stands for $(a \lesssim b) \wedge (b \lesssim a)$. If we need to allow the implicit constant to depend on parameters we indicate this by subscripts, thus $a \lesssim_p b$ stands for $a \leqslant c_p b$ where c_p is allowed to depend only on p, and similarly for the notations $\gtrsim_p$ and $\asymp_p$.

We conjecture that when $p \in (2,\infty)$ the lower bound in Theorem (2) could be improved to

$$c_{(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)}(X_n, d_{X_n}^\alpha) \gtrsim_p (\log n)^{\alpha-\frac{1}{2}},$$

and, correspondingly, that the conclusion of Theorem 1 could be improved to state that if $p \in (2,\infty)$ then $c_{(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)}\big(X, \sqrt{d_X}\big) \lesssim_p 1$ for every finite metric space (X, d_X); see Question 23 below.

There are several motivations for our investigations that led to Theorem 1 and Theorem 2. Notably, we are inspired by a longstanding open question of Bourgain [13], as well as fundamental questions on the geometry of Alexandrov spaces. We shall now explain these links.

1.3. **Alexandrov geometry.** We need to briefly present some standard background on metric spaces that are either nonnegatively curved or nonpositively curved in the sense of Alexandrov; the relevant background can be found in e.g. [18, 15]. Let (X, d_X) be a complete geodesic metric space. Recall that $w \in X$ is called a metric midpoint of $x, y \in X$ if $d_X(x, w) = d_X(y, w) = d_X(x, y)/2$. The metric space (X, d_X) is said to be an Alexandrov space of nonnegative curvature if for every $x, y, z \in X$ and every metric midpoint w of x, y,

$$d_X(x, y)^2 + 4d_X(z, w)^2 \geqslant 2d_X(x, z)^2 + 2d_X(y, z)^2. \tag{1}$$

Correspondingly, the metric space (X, d_X) is said to be an Alexandrov space of nonpositive curvature, or a Hadamard space, if for every $x, y, z \in X$ and every metric midpoint w of x, y,

$$d_X(x, y)^2 + 4d_X(z, w)^2 \leqslant 2d_X(x, z)^2 + 2d_X(y, z)^2. \tag{2}$$

If (X,d_X) is a Hilbert space then, by the parallelogram identity, the inequalities (1) and (2) hold true as equalities (with $w=(x+y)/2$). So, (1) and (2) are both natural relaxations of a stringent Hilbertian identity (both relaxations have far-reaching implications). A complete Riemannian manifold is an Alexandrov space of nonnegative curvature if and only if its sectional curvature is nonnegative everywhere, and a complete simply connected Riemannian manifold is a Hadamard space if and only if its sectional curvature is nonpositive everywhere.

Following [78], it was shown in [89, Proposition 2.10] and [52, Appendix A] that $\mathcal{P}_2(\mathbb{R}^n)$ is an Alexandrov space of nonnegative curvature for every $n\in\mathbb{N}$; more generally, if (X,d_X) is an Alexandrov space of nonnegative curvature then so is $\mathcal{P}_2(X)$. It therefore follows from Theorem 1 that there exists an Alexandrov space (Y,d_Y) of nonnegative curvature that contains a bi-Lipschitz copy of the 1/2-snowflake of every finite metric space, with distortion at most $1+\varepsilon$ for every $\varepsilon>0$. When this happens, we shall say that (Y,d_Y) is 1/2-snowflake universal.

1.4. **Subsets of Alexandrov spaces.** It is a longstanding open problem, stated by Gromov in [31, Section 1.19_+] and [32, §15(b)], as well as in, say, [25, 1, 90], to find an intrinsic characterization of those metric spaces that admit a bi-Lipschitz, or even isometric, embedding into an Alexandrov space of either nonnegative or nonpositive curvature.

Berg and Nikolaev [8, 9] (see also [85]) proved that a complete metric space (X,d_X) is a Hadamard space if and only if it is geodesic and every $x_1,x_2,x_3,x_4\in X$ satisfy

$$d_X(x_1,x_3)^2+d_X(x_2,x_4)^2\leqslant d_X(x_1,x_2)^2+d_X(x_2,x_3)^2+d_X(x_3,x_4)^2+d_X(x_4,x_1)^2. \qquad (3)$$

Inequality (3) is known in the literature under several names, including Enflo's "roundness 2 property" (see [23]), "the short diagonal inequality" (see [54]), or simply "the quadrilateral inequality," and it has a variety of important applications. Another characterization of this nature is due to Foertsch, Lytchak and Schroeder [25], who proved that a complete metric space (X,d_X) is a Hadamard space if and only if it is geodesic, every $x_1,x_2,x_3,x_4\in X$ satisfy the inequality

$$d_X(x_1,x_3)\cdot d_X(x_2,x_4)\leqslant d_X(x_1,x_2)\cdot d_X(x_3,x_4)+d_X(x_2,x_3)\cdot d_X(x_1,x_4), \qquad (4)$$

and if w is a metric midpoint of x_1 and x_2 and z is a metric midpoint of x_3 and x_4 then we have

$$d_X(w,z)\leqslant \frac{d_X(x_1,x_3)+d_X(x_2,x_4)}{2}. \qquad (5)$$

(4) is called the Ptolemy inequality [26], and condition (5) is called Busemann convexity [19].

Turning now to characterizations of nonnegative curvature, Lebedeva and Petrunin [47] proved that a complete metric space (X,d_X) is an Alexandrov space of nonnegative curvature if and only if it is geodesic and every $x,y,z,w\in X$ satisfy

$$d_X(x,w)^2+d_X(y,w)^2+d_X(z,w)^2\geqslant \frac{d_X(x,y)^2+d_X(x,z)^2+d_X(y,z)^2}{3}.$$

Another (related) important characterization of Alexandrov spaces of nonnegative curvature asserts that a metric space (X,d_X) is an Alexandrov spaces of nonnegative curvature if and only if it is geodesic and for every finitely supported X-valued random variable Z we have

$$\mathbb{E}\big[d_X(Z,Z')^2\big]\leqslant 2\inf_{x\in X}\mathbb{E}\big[d_X(Z,x)^2\big], \qquad (6)$$

where Z' is an independent copy of Z. The above characterization is due to Sturm [87], with the fact that nonnegative curvature in the sense of Alexandrov implies the validity of (6) being due to Lang and Schroeder [46]. Following e.g. [97], condition (6) (which we shall use in Section 3) is therefore called the Lang–Schroeder–Sturm inequality.

The above statements are interesting characterizations of spaces that are isometric to Alexandrov spaces of either nonpositive or nonnegative curvature, but they fail to characterize *subsets* of such spaces, since they require additional convexity properties of the metric space in question, such as being geodesic or Busemann convex. These assumptions are not intrinsic because they stipulate the existence of auxiliary points (metric midpoints) which may fall outside the given subset. Furthermore, these characterizations are *isometric* in nature, thus failing to address the important question of understanding when, given $D \in (1,\infty)$, a metric space (X,d_X) embeds with distortion at most D into some Alexandrov space of either nonpositive or nonnegative curvature. One can search for such characterizations only among families of *quadratic metric inequalities*, as we shall now explain; in our context this is especially natural because the definitions (1) and (2) are themselves quadratic.

1.4.1. *Quadratic metric inequalities.* For $n \in \mathbb{N}$ and n by n matrices $A=(a_{ij}), B=(b_{ij}) \in M_n(\mathbb{R})$ with nonnegative entries, say that a metric space (X,d_X) satisfies the (A,B)-quadratic metric inequality if for every $x_1,\ldots,x_n \in X$ we have

$$\sum_{i=1}^n\sum_{j=1}^n a_{ij} d_X(x_i,x_j)^2 \leqslant \sum_{i=1}^n\sum_{j=1}^n b_{ij} d_X(x_i,x_j)^2.$$

The property of satisfying the (A,B)-quadratic metric inequality is clearly preserved by forming Pythagorean products, i.e., if (X,d_X) and (Y,d_Y) both satisfy the (A,B)-quadratic metric inequality then so does their Pythagorean product $(X\oplus Y)_2$. Here $(X\oplus Y)_2$ denotes the space $X\times Y$, equipped with the metric that is defined by

$$\forall (a,b),(\alpha,\beta) \in X\times Y, \qquad d_{(X\oplus Y)_2}\big((a,b),(\alpha,\beta)\big) \stackrel{\mathrm{def}}{=} \sqrt{d_X(a,\alpha)^2+d_Y(b,\beta)^2}.$$

The (A,B)-quadratic metric inequality is also preserved by ultraproducts (see e.g. [39, Section 2.4] for background on ultraproducts of metric spaces), and it is a bi-Lipschitz invariant in the sense that if (X,d_X) embeds with distortion at most $D\in[1,\infty)$ into (Y,d_Y), and (Y,d_Y) satisfies the (A,B)-quadratic metric inequality then (X,d_X) satisfies the (A,D^2B)-quadratic metric inequality.

The following proposition is a converse to the above discussion.

Proposition 3. *Let $\mathcal{F}$ be a family of metric spaces that is closed under dilation and Pythagorean products, i.e., if $(U,d_U),(V,d_V)\in\mathcal{F}$ and $s\in(0,\infty)$ then also $(U,sd_U)\in\mathcal{F}$ and $(U\oplus V)_2\in\mathcal{F}$. Fix $D\in[1,\infty)$ and $n\in\mathbb{N}$. Then an n-point metric space (X,d_X) satisfies*

$$\inf_{(Y,d_Y)\in\mathcal{F}} c_Y(X) \leqslant D$$

if and only if for every two n by n matrices $A,B\in M_n(\mathbb{R})$ with nonnegative entries such that every $(Z,d_Z)\in\mathcal{F}$ satisfies the (A,B)-quadratic metric inequality, we also have that (X,d_X) satisfies the (A,D^2B)-quadratic metric inequality.

The proof of Proposition 3 appears in Section 4 below and consists of a duality argument that mimics the proof of Proposition 15.5.2 in [54], which deals with embeddings into Hilbert space.

Remark 4. It is a formal consequence of Proposition 3 that if the family of metric spaces $\mathcal{F}$ is also closed under ultraproducts, as are Alexandrov spaces with upper or lower curvature bounds (see e.g. [39, Section 2.4]), then one does not need to restrict to finite metric spaces. Namely, in this case a metric space (X,d_X) admits a bi-Lipschitz embedding into some $(Y,d_Y)\in\mathcal{F}$ if and only if there exists $D\in[1,\infty)$ such that (X,d_X) satisfies the (A,D^2B)-quadratic metric inequality for every two n by n matrices $A,B\in M_n(\mathbb{R})$ with nonnegative entries such that every $(Z,d_Z)\in\mathcal{F}$ satisfies the (A,B)-quadratic metric inequality.

Remark 5. The Ptolemy inequality (4) is not a quadratic metric inequality, yet it holds true in any Hadamard space. Proposition 3 implies that the Ptolemy inequality could be deduced from quadratic metric inequalities that hold true in Hadamard spaces. This is carried out explicitly in Section 5 below, yielding an instructive proof (and strengthening) of the Ptolemy inequality in Hadamard spaces that is conceptually different from its previously known proofs [25, 17].

Theorem 1 implies that all the quadratic metric inequalities that hold true in every Alexandrov space of nonnegative curvature "trivialize" if one does not square the distances. Specifically, since $\mathcal{P}_2(\mathbb{R}^3)$ is an Alexandrov space of nonnegative curvature, the following statement is an immediate consequence of Theorem 1.

Theorem 6. *Suppose that $A, B \in M_n(\mathbb{R})$ are n by n matrices with nonnegative entries such that every Alexandrov space of nonnegative curvature satisfies the (A,B)-quadratic metric inequality. Then for every metric space (X, d_X) and every $x_1, \ldots, x_n \in X$ we have*

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} d_X(x_i, x_j) \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} d_X(x_i, x_j). \tag{7}$$

While Theorem 6 does not answer the question of characterizing those quadratic metric inequalities that hold true in any Alexandrov space of nonnegative curvature, it does show that such inequalities rely crucially on the fact that distances are being squared, i.e., if one removes the squares then one arrives at an inequality (7) which must be nothing more than a consequence of the triangle inequality.

Obtaining a full characterization of those quadratic metric inequalities that hold true in any Alexandrov space of nonnegative curvature remains an important challenge. Many such inequalities are known, including, as shown by Ohta [76], Markov type 2 (note, however, that the supremum of the Markov type 2 constants of all Alexandrov spaces of nonnegative curvature is an unknown universal constant [77]; we obtain the best known bound on this constant in Corollary 26 below). Another family of nontrivial quadratic metric inequalities that hold true in any Alexandrov space of nonnegative curvature is obtained in [3], where it is shown that all such spaces have Markov convexity 2. By these observations combined with the nonlinear Maurey–Pisier theorem [57], we know that there exists $q < \infty$ such that any Alexandrov space of nonnegative curvature has metric cotype q. It is natural to conjecture that one could take $q = 2$ here, but at present this remains open. For more on the notions discussed above, i.e., Markov type, Markov convexity and metric cotype, as well as their applications, see the survey [66] and the references therein.

The above discussion in the context of Hadamard spaces remains an important open problem. At present we do not know of any metric space (X, d_X) such that the metric space $(X, \sqrt{d_X})$ fails to admit a bi-Lipschitz embedding into some Hadamard space. More generally, while a variety of nontrivial quadratic metric inequalities are known to hold true in any Hadamard space, a full characterization of such inequalities remains elusive. In Section 5 below we formulate a systematic way to generate such inequalities, posing the question whether the hierarchy of inequalities thus obtained yields a characterization of those metric spaces that admit a bi-Lipschitz embedding into some Hadamard space.

1.4.2. *Uniform, coarse and quasisymmetric embeddings.* A metric space (X, d_X) is said to embed uniformly into a metric space (Y, d_Y) if there exists an injection $f : X \to Y$ such that both f and f^{-1} are uniformly continuous. (X, d_X) is said [30] to embed coarsely into (Y, d_Y) if there exists $f : X \to Y$ and nondecreasing functions $\alpha, \beta : [0, \infty) \to [0, \infty)$ with $\lim_{t\to\infty} \alpha(t) = \infty$ such that

$$\forall\, x, y \in X, \qquad \alpha(d_X(x, y)) \leqslant d_Y(f(x), f(y)) \leqslant \beta(d_X(x, y)). \tag{8}$$

(X, d_X) is said [10, 92] to admit a quasisymmetric embedding into (Y, d_Y) if there exists an injection $f : X \to Y$ and $\eta : (0, \infty) \to (0, \infty)$ with $\lim_{t\to 0} \eta(t) = 0$ such that for every distinct $x, y, z \in X$,

$$\frac{d_Y(f(x), f(y))}{d_Y(f(x), f(z))} \leqslant \eta\left(\frac{d_X(x, y)}{d_X(x, z)}\right).$$

A direct combination of Theorem 1 with the results of [57, 65] shows that $\mathcal{P}_2(\mathbb{R}^3)$ does not embed even in the above weak senses into any Banach space of nontrivial (Rademacher) type; we refer to the survey [55] and the references therein for more on the notion of type of Banach spaces. In particular, $\mathcal{P}_2(\mathbb{R}^3)$ fails to admit such embeddings into any $L_p(\mu)$ space for finite p (for the case $p = 1$, use the fact that the 1/2-snowflake of an $L_1(\mu)$ space embeds isometrically into a Hilbert space; see [95]), or, say, into any uniformly convex Banach space. It remains an interesting open question whether or not these assertions also hold true for $\mathcal{P}_2(\mathbb{R}^2)$.

Theorem 7. *If $p > 1$ then $\mathcal{P}_p(\mathbb{R}^3)$ does not admit a uniform, coarse or quasisymmetric embedding into any Banach space of nontrivial type.*

Note that a positive resolution of a key conjecture of [57], namely the first question in Section 8 of [57], would "upgrade" Theorem 7 to the (best possible) assertion that $\mathcal{P}_2(\mathbb{R}^3)$ does not admit a uniform, coarse or quasisymmetric embedding into any Banach space of finite cotype.

Remark 8. Very few other examples of Alexandrov spaces of nonnegative curvature with poor embeddability properties into Banach spaces are known, all of which are not known to satisfy properties as strong as the conclusion of Theorem 7. Specifically, in [3] it is shown that $\mathcal{P}_2(\mathbb{R}^2)$ fails to admit a bi-Lipschitz embedding into L_1. A construction with stronger properties follows from the earlier work [38], combined with the recent methods of [67]. Specifically, it follows from [38] and [67] that for every $n \in \mathbb{N}$ there exists a lattice $\Lambda_n \subseteq \mathbb{R}^n$ of rank n such that if we consider the following infinite Pythagorean product of flat tori

$$\mathcal{T} \stackrel{\mathrm{def}}{=} \Big(\bigoplus_{n=1}^{\infty} \mathbb{R}^n/\Lambda_n \Big)_2, \tag{9}$$

then $\mathcal{T}$ fails to admit a uniform or coarse embedding into a certain class of Banach spaces that includes all Banach lattices of finite cotype and all the noncommutative L_p spaces for finite $p \geqslant 1$. Since for every $n \in \mathbb{N}$ the sectional curvature of $\mathbb{R}^n/\Lambda_n$ vanishes, it is an Alexandrov space of nonnegative curvature, and therefore so is the Pythagorean product $\mathcal{T}$. It remains an interesting open question whether or not $\mathcal{T}$ admits a uniform, coarse or quasisymmetric embedding into some Banach space of nontrivial type, and, for that matter, even whether or not $\mathcal{T}$ is 1/2-snowflake universal. We speculate that the answer to the latter question is negative.

1.4.3. *Expanders with respect to Alexandrov spaces.* Fixing an integer $k \geqslant 3$, an unbounded sequence of k-regular finite graphs $\{(V_j, E_j)\}_{j=1}^{\infty}$ is said to be an expander with respect to a metric space (X, d_X) if for every $j \in \mathbb{N}$ and $\{x_u\}_{u\in V_j} \subseteq X$ we have

$$\frac{1}{|V_j|^2} \sum_{(u,v)\in V_j\times V_j} d_X(x_u, x_v)^2 \asymp_X \frac{1}{|E_j|} \sum_{\{u,v\}\in E_j} d_X(x_u, x_v)^2. \tag{10}$$

Unless X is a singleton, a sequence of expanders with respect to (X, d_X) must also be a sequence of expanders in the classical (combinatorial) sense. See [73, 61, 62, 67, 75] and the references therein for background on expanders with respect to metric spaces and their applications.

In contrast to the case of classical expanders, the question of understanding when a metric space X admits an expander sequence seems to be very difficult (even in the special case when X is a Banach space), with limited availability of methods [53, 79, 42, 43, 61, 50, 67, 62, 63] for establishing metric inequalities such as (10). Theorem 1 implies that $\mathcal{P}_p(\mathbb{R}^3)$ fails to admit a sequence of expanders for every $p \in (1, \infty)$. The particular case $p = 2$ establishes for the first time the (arguably surprising) fact that there exists an Alexandrov space of nonnegative curvature with respect to which expanders do not exist.

Theorem 9. *For $p > 1$ no sequence of bounded degree graphs is an expander with respect to $\mathcal{P}_p(\mathbb{R}^3)$.*

To deduce Theorem 9 from Theorem 1, use an argument of Gromov [33] (which is reproduced in [61, Section 1.1]), to deduce that if $\{G_n = (V_n, E_n)\}_{n=1}^{\infty}$ were a k-regular expander with respect to $\mathcal{P}_p(\mathbb{R}^3)$ then, denoting the shortest-path metric that G_n induces on V_n by d_n (the assumption that G_n is an expander with respect to a non-singleton metric space implies that it is a classical expander, hence connected), the metric spaces $\{(V_n, d_n)\}_{n=1}^{\infty}$ fail to admit a coarse embedding into $\mathcal{P}_p(\mathbb{R}^3)$ with any moduli $\alpha, \beta : [0,\infty) \to [0,\infty)$ as in (8) that are independent of n. This contradicts the fact that by Theorem 1 we know that for every $n \in \mathbb{N}$ the finite metric space (V_n, d_n) embeds coarsely into $\mathcal{P}_p(\mathbb{R}^3)$ with moduli $\alpha(t) = t^{1/p}$ and, say, $\beta(t) = 2t^{1/p}$.

The above question for Hadamard spaces remains an important open problem which goes back at least to [33, 73]. See [62] for more on this theme, where it is shown that there exists a Hadamard space with respect to which random regular graphs are asymptotically almost surely not expanders. We also ask whether or not the Alexandrov space of nonnegative curvature $\mathcal{T}$ of Remark 8 admits a sequence of bounded degree expanders; we speculate that it does.

1.5. **The universality problem for $\mathcal{P}_1(\mathbb{R}^k)$.** A metric space (Y, d_Y) is said to be (finitely) universal if there exists $K \in (0,\infty)$ such that $c_Y(X) \leqslant K$ for *every finite* metric space (X, d_X).

In [13] Bourgain asked whether $(\mathcal{P}_1(\mathbb{R}^2), \mathsf{W}_1)$ is not universal. He actually formulated this question as asking whether a certain Banach space (namely, the dual of the Lipschitz functions on the square $[0,1]^2$), which we denote for the sake of the present discussion by Z, has finite Rademacher cotype, but this is equivalent to the above formulation in terms of the universality of $(\mathcal{P}_1(\mathbb{R}^2), \mathsf{W}_1)$. It is not necessary to be familiar with the notion of cotype in order to understand the ensuing discussion, so readers can consider only the above formulation of Bourgain's question. However, for experts we shall now briefly justify this equivalence. For Banach spaces the property of not being universal is equivalent to having finite Rademacher cotype, as follows from Ribe's theorem [**?**] and the Maurey–Pisier theorem [**?**]. As explained in [72], every finite subset of Z embeds into $\mathcal{P}_1(\mathbb{R}^2)$ with distortion arbitrarily close to 1, and, conversely, every finite subset of $\mathcal{P}_1(\mathbb{R}^2)$ embeds into Z with distortion arbitrarily close to 1. Hence Z is universal if and only if $\mathcal{P}_1(\mathbb{R}^2)$ is universal. So, Z has finite Rademacher cotype if and only if $\mathcal{P}_1(\mathbb{R}^2)$ is not universal.

Bourgain proved in [13] that $(\mathcal{P}_1(\ell_1), \mathsf{W}_1)$ is universal (despite the fact that ℓ_1 is not universal), but it remains an intriguing open question to determine whether or not $(\mathcal{P}_1(\mathbb{R}^k), \mathsf{W}_1)$ is universal for any finite $k \in \mathbb{N}$, the case $k = 2$ being most challenging. Here we show that Wasserstein spaces do exhibit some universality phenomenon even when the underlying metric space is a finite dimensional Euclidean space, but we fall short of addressing the universality problem for $\mathcal{P}_1(\mathbb{R}^k)$. Specifically, Theorem 1 asserts that $(\mathcal{P}_p(\mathbb{R}^3), \mathsf{W}_p)$ is universal with respect to $1/p$-snowflakes of metric spaces, and if $p \in (1,2]$ then this cannot be improved to α-snowflakes for any $\alpha > 1/p$, by Theorem (2). The $1/p$-snowflake of (X, d_X) becomes "closer" to (X, d_X) itself as $p \to 1$, and at the same time $(\mathcal{P}_p(\mathbb{R}^3), \mathsf{W}_p)$ becomes "closer" to $(\mathcal{P}_1(\mathbb{R}^3), \mathsf{W}_1)$, but Theorem 1 fails to imply the universality of $(\mathcal{P}_1(\mathbb{R}^3), \mathsf{W}_1)$ because the embeddings that we construct in Theorem 1 degenerate as $p \to 1$.

Remark 10. The universality problem for $\mathcal{P}_1(\mathbb{R}^k)$ belongs to longstanding traditions in functional analysis. As Bourgain explains in [13], one motivation for his question is an idea of W. B. Johnson to "linearize" bi-Lipschitz classification problems by examining the geometry of the corresponding Banach spaces of Lipschitz functions defined on the metric spaces in question. For this "functorial linearization" to succeed, one needs to sufficiently understand the linear structure of the spaces of Lipschitz functions on metric spaces, but unfortunately these are wild spaces that are poorly understood. The universality problem for $\mathcal{P}_1(\mathbb{R}^k)$ highlights this situation by asking a basic geometric question (universality) about the dual of the space of Lipschitz functions on $\mathbb{R}^k$. Despite these difficulties, in recent years the above approach to bi-Lipschitz classification problems has been successfully developed, notably by Godefroy and Kalton [28] who, among other results, deduced from this approach that the Bounded Approximation Property (BAP) is preserved under bi-Lipschitz homeomorphisms of Banach spaces. In addition to being motivated by potential applications, the universality problem for $\mathcal{P}_1(\mathbb{R}^k)$ relates to old questions on the structure of classical function spaces: here the spaces in question are the Lipschitz functions on $\mathbb{R}^k$, which are closely related to the spaces $C^1(\mathbb{R}^k)$ whose linear structure (in particular its dependence on k) remains a major mystery that goes back to Banach's seminal work. Understanding the universality of classical Banach spaces and their duals has attracted many efforts over the past decades, notable examples of which include work [80, 11] on the (non)universality of the dual of the Hardy space $H^\infty(S^1)$, work [93, 81, 40, 14] on the universality of the span in $C(G)$ of a subset of characters of a compact Abelian group G, and work [91, 82, 16] on the universality of projective tensor products. Despite these efforts, understanding the universality of $\mathcal{P}_1(\mathbb{R}^k)$ (equivalently, whether or not the dual of the space of Lipschitz functions on $\mathbb{R}^k$ has finite cotype) remains a remarkably stubborn open problem.

Our proof of Theorem 1 relies on the fact that the underlying Euclidean space is (at least) 3-dimensional, so it remains open whether or not, say, the 1/2-snowflake of every finite metric space embeds with $O(1)$ distortion into $(\mathcal{P}_2(\mathbb{R}^2), \mathsf{W}_2)$. In [3] it is proved that every finite subset of the metric space $(\mathcal{P}_1(\mathbb{R}^2), \sqrt{\mathsf{W}_1})$, i.e., the 1/2-snowflake of $(\mathcal{P}_1(\mathbb{R}^2), \mathsf{W}_1)$, embeds with $O(1)$ distortion into $(\mathcal{P}_2(\mathbb{R}^2), \mathsf{W}_2)$. Thus, if $(\mathcal{P}_1(\mathbb{R}^2), \mathsf{W}_1)$ were universal (i.e., if the universality problem for $\mathcal{P}_1(\mathbb{R}^2)$ had a negative answer) then it would follow that the 1/2-snowflake of every finite metric space embeds with $O(1)$ distortion into $(\mathcal{P}_2(\mathbb{R}^2), \mathsf{W}_2)$.

Remark 11. Another interesting open question is whether or not $\mathcal{P}_1(\mathbb{R}^3)$ (or $\mathcal{P}_1(\mathbb{R}^2)$ for that matter) is 1/2-snowflake universal. There is a perceived analogy between the spaces $\mathcal{P}_p(X)$ and $L_p(\mu)$ spaces, with the spaces $\mathcal{P}_p(X)$ sometimes being referred to as the geometric measure theory analogues of $L_p(\mu)$ spaces. It would be very interesting to investigate whether or not this analogy could be put on firm footing. As an example of a concrete question along these lines, since L_2 is isometric to a subspace of L_p, we ask for a characterization of those metric spaces X for which $\mathcal{P}_2(X)$ admits a bi-Lipschitz embedding into $\mathcal{P}_p(X)$, or, less ambitiously, when does there exist $D(X) \in [1, \infty)$ such that every finite subset of $\mathcal{P}_2(X)$ embeds into $\mathcal{P}_p(X)$ with distortion $D(X)$. If this were true when $X = \mathbb{R}^3$ or $X = \mathbb{R}^2$ (it is easily seen to be true when $X = \mathbb{R}$) and $p = 1$ then it would follow from Theorem 1 that $\mathcal{P}_1(\mathbb{R}^3)$ (respectively $\mathcal{P}_1(\mathbb{R}^2)$) is 1/2-snowflake universal. By [57], this, in turn, would imply that $\mathcal{P}_1(\mathbb{R}^3)$ (respectively $\mathcal{P}_1(\mathbb{R}^2)$) fails to admit a coarse, uniform or quasisymmetric embedding into L_1, thus strengthening results of [72] via an approach that is entirely different from that of [72]. There are many additional open questions that follow from the analogy between Wasserstein p spaces and $L_p(\mu)$ spaces, including various questions about the evaluation of the metric type and cotype of $\mathcal{P}_p(X)$; see Question 23 below for more on this interesting research direction.

1.5.1. *Towards the metric cotype dichotomy problem.* The following theorem was proved in [57]; see [56, 59, 60] for more information on metric dichotomies of this type.

Theorem 12 (Metric cotype dichotomy [57])**.** *Let (X, d_X) be a metric space that isn't universal. There exists $\alpha(X) \in (0, \infty)$ and finite metric spaces $\{(M_n, d_{M_n})\}_{n=1}^\infty$ with $\lim_{n\to\infty} |M_n| = \infty$ and*

$$\forall\, n \in \mathbb{N}, \qquad c_X(M_n) \geqslant (\log |M_n|)^{\alpha(X)}.$$

A central question that was left open in [57], called the *metric cotype dichotomy problem*, is whether the exponent $\alpha(X) \in (0, \infty)$ of Theorem 12 can be taken to be a universal constant, i.e.,

Question 13 (Metric cotype dichotomy problem [57])**.** Does there exist $\alpha \in (0, 1]$ such that *every* non-universal metric space X admits a sequence of finite metric spaces $\{(M_n, d_{M_n})\}_{n=1}^\infty$ with $\lim_{n\to\infty} |M_n| = \infty$ that satisfies $c_X(M_n) \geqslant (\log |M_n|)^\alpha$?

It is even unknown whether or not in Question 13 one could take $\alpha = 1$ (by Bourgain's embedding theorem [12], the best one could hope for here is $\alpha = 1$). A positive answer to the following question would resolve the metric cotype dichotomy problem negatively; this question corresponds to asking if Theorem 2 is sharp when $p \in (1,2]$ and $\alpha = 1$ (the same question when $\alpha \in (1/p, 1)$ is also open).

Question 14. Is it true that for $p \in (1,2]$ and $n \in \mathbb{N}$ every n-point metric space (X, d_X) satisfies

$$c_{\mathcal{P}_p(\mathbb{R}^3)}(X) \lesssim_p (\log n)^{1-\frac{1}{p}}?$$

A positive answer to Question (14) would imply that $\alpha(\mathcal{P}_p(\mathbb{R}^3)) \leqslant 1 - 1/p$, using the notation of Theorem 12. Taking $p \to 1^+$, it would therefore follow that there is no $\alpha > 0$ as in Question 13.

We believe that Question 14 is an especially intriguing challenge in embedding theory (for a concrete and natural target space) because a positive answer, in addition to resolving the metric cotype dichotomy problem, would require an interesting new construction, and a negative answer would require devising a new bi-Lipschitz invariant that would serve as an obstruction for embeddings into Wasserstein spaces.

Focusing for concreteness on the case $p = 2$, Question 14 asks whether $c_{\mathcal{P}_2(\mathbb{R}^3)}(X) \lesssim \sqrt{\log n}$ for every n-point metric space (X, d_X). Note that Theorem 1 implies that (X, d_X) embeds into $\mathcal{P}_2(\mathbb{R}^3)$ with distortion at most the square root of the *aspect ratio* of (X, d_X), i.e.,

$$c_{(\mathcal{P}_2(\mathbb{R}^3), \mathsf{W}_2)}(X, d_X) \leqslant \sqrt{\frac{\mathrm{diam}(X, d_X)}{\min_{\substack{x,y \in X\\ x \neq y}} d_X(x,y)}}, \tag{11}$$

but we are asking here for the largest possible growth rate of the distortion of X into $\mathcal{P}_2(\mathbb{R}^3)$ in terms of the cardinality of X. While for certain embedding results there are standard methods (see e.g. [6, 34, 58]) for replacing the dependence on the aspect ratio of a finite metric space by a dependence on its cardinality, these methods do not seem to apply to our embedding in (11). See Section 6 below for further discussion.

2. Proof of Theorem 1

In what follows fix $n \in \mathbb{N}$ and an n-point metric space (X, d_X). Write $X = \{x_1, \ldots, x_n\}$ and fix $\phi : \{1, \ldots, n\} \times \{1, \ldots, n\} \to \{1, \ldots, n^2\}$ to be an arbitrary bijection between $\{1, \ldots, n\} \times \{1, \ldots, n\}$ and $\{1, \ldots, n^2\}$. Below it will be convenient to use the following notation.

$$m \stackrel{\mathrm{def}}{=} \min_{\substack{x,y \in X\\ x \neq y}} d_X(x,y)^{\frac{1}{p}} \qquad \text{and} \qquad M \stackrel{\mathrm{def}}{=} \max_{x,y \in X} d_X(x,y)^{\frac{1}{p}}. \tag{12}$$

Fix $K \in \mathbb{N}$. Denoting the standard basis of $\mathbb{R}^3$ by $e_1 = (1,0,0)$, $e_2 = (0,1,0)$, $e_3 = (0,0,1)$, for every $i, j \in \{1, \ldots, n\}$ with $i < j$ define five families of points in $\mathbb{R}^3$ by setting for $s \in \{0, \ldots, K\}$,

$$Q_s^1(i,j) \stackrel{\mathrm{def}}{=} \frac{Mi}{m} e_1 + \frac{M\phi(i,j)s}{mK} e_2, \tag{13}$$

$$Q_s^2(i,j) \stackrel{\mathrm{def}}{=} \frac{Mi}{m} e_1 + \frac{M\phi(i,j)}{m} e_2 + \frac{Ms}{mK} e_3, \tag{14}$$

$$Q_s^3(i,j) \stackrel{\mathrm{def}}{=} \frac{M(s(j-i) + Ki) + (K-s) d_X(x_i, x_j)^{\frac{1}{p}}}{mK} e_1 + \frac{M\phi(i,j)}{m} e_2 + \frac{M}{m} e_3, \tag{15}$$

$$Q_s^4(i,j) \stackrel{\mathrm{def}}{=} \frac{Mj}{m} e_1 + \frac{M\phi(i,j)}{m} e_2 + \frac{M(K-s)}{mK} e_3, \tag{16}$$

$$Q_s^5(i,j) \stackrel{\mathrm{def}}{=} \frac{Mj}{m} e_1 + \frac{M(K-s)\phi(i,j)}{mK} e_2. \tag{17}$$

Then $Q_K^1(i,j) = Q_0^2(i,j)$, $Q_K^3(i,j) = Q_0^4(i,j)$ and $Q_K^4(i,j) = Q_0^5(i,j)$, so the total number of points thus obtained equals $5(K+1) - 3 = 5K + 2$.

Define $\mathcal{B} \subseteq \mathbb{R}^3$ by setting

$$\mathcal{B} \stackrel{\mathrm{def}}{=} \bigcup_{\substack{i,j \in \{1,\ldots,n\} \\ i<j}} \mathcal{B}_{ij}, \tag{18}$$

where for every $i,j \in \{1,\ldots,n\}$ with $i < j$ we write

$$\mathcal{B}_{ij} \stackrel{\mathrm{def}}{=} \bigcup_{s=0}^{K} \left\{Q_s^1(i,j), Q_s^2(i,j), Q_s^3(i,j), Q_s^4(i,j), Q_s^5(i,j)\right\}. \tag{19}$$

Hence $|\mathcal{B}_{ij}| = 5K + 2$. We also define $\mathcal{C} \subseteq \mathbb{R}^3$ by

$$\mathcal{C} \stackrel{\mathrm{def}}{=} \mathcal{B} \smallsetminus \left\{\frac{Mi}{m} e_1 : \ i \in \{1,\ldots,n\}\right\}. \tag{20}$$

Note that by (13) we have $(Mi/m)e_1 = Q_0^1(i,j)$ if $i,j \in \{1,\ldots,n\}$ satisfy $i < j$, and by (17) we have $(Mi/m)e_1 = Q_K^5(\ell,i)$ if $\ell, i \in \{1,\ldots,n\}$ satisfy $\ell < i$. Thus $\mathcal{C}$ corresponds to removing from $\mathcal{B}$ those points that lie on the x-axis. In what follows, we denote $N = |\mathcal{C}| + 1$. Finally, for every $i \in \{1,\ldots,n\}$ we define $\mathcal{C}_i \subseteq \mathbb{R}^3$ by

$$\mathcal{C}_i \stackrel{\mathrm{def}}{=} \mathcal{C} \cup \left\{\frac{Mi}{m} e_1\right\}. \tag{21}$$

Hence $|\mathcal{C}_i| = N$. Our embedding $f : X \to \mathcal{P}_p(\mathbb{R}^3)$ will be given by

$$\forall\, j \in \{1,\ldots,n\}, \qquad f(x_j) \stackrel{\mathrm{def}}{=} \frac{1}{N} \sum_{u \in \mathcal{C}_j} \delta_u, \tag{22}$$

where, as usual, δ_u is the point mass at u. Thus $f(x_j)$ is the uniform probability measure over $\mathcal{C}_j$. A schematic depiction of the above construction appears in Figure 1 below.

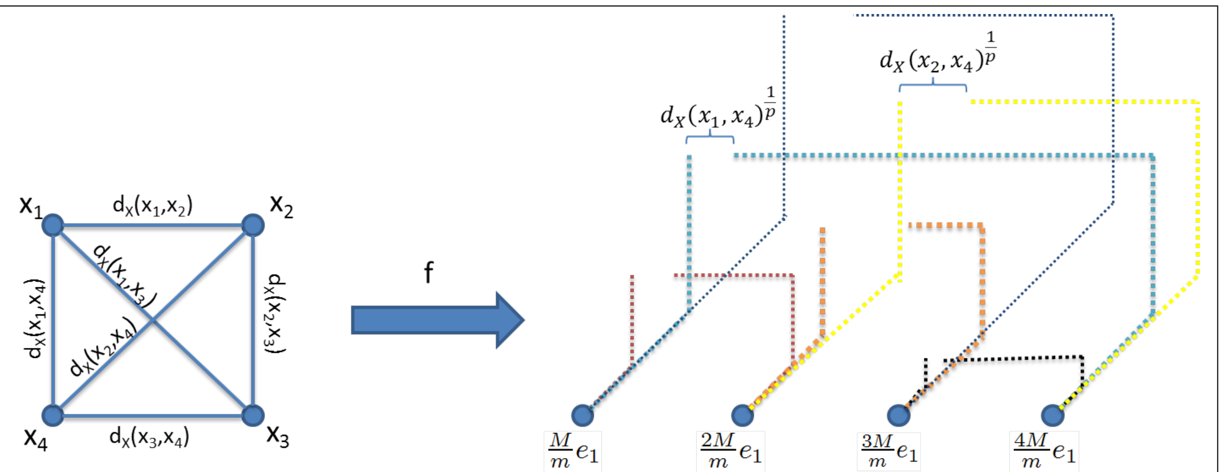

FIGURE 1. A schematic depiction of the embedding $f : X \to \mathcal{P}_p(\mathbb{R}^3)$ for a four-point metric space $(X, d_X) = (\{x_1, x_2, x_3, x_4\}, d_X)$. Here the x-axis is the horizontal direction, the z-axis is the vertical direction and the y-axis is perpendicular to the page plane. Recall that m and M are defined in (12).

Lemma 15 below estimates the distortion of f, proving Theorem 1.

Lemma 15. *Fix $\varepsilon \in (0,1)$ and $p \in (1,\infty)$. Let $f : X \to \mathcal{P}_p(\mathbb{R}^3)$ be the mapping appearing in (22), considered as a mapping from the snowflaked metric space $(X, d_X^{1/p})$ to the metric space $(\mathcal{P}_p(\mathbb{R}^3), \mathsf{W}_p)$. Then, recalling the definitions of m and M in (12), we have*

$$K \geqslant \left(\frac{5M^p n^{2p}}{pm^p\varepsilon}\right)^{\frac{1}{p-1}} \implies \mathbf{dist}(f) \leqslant 1+\varepsilon. \tag{23}$$

Proof. We shall show that under the assumption on K that appears in (23) we have

$$\forall\, i,j \in \{1,\ldots,n\}, \qquad \left(\frac{d_X(x_i,x_j)}{m^pN}\right)^{\frac{1}{p}} \leqslant \mathsf{W}_p(f(x_i),f(x_j)) \leqslant (1+\varepsilon)\left(\frac{d_X(x_i,x_j)}{m^pN}\right)^{\frac{1}{p}}, \tag{24}$$

where we recall that we defined N to be equal to $|\mathcal{C}|+1$ for $\mathcal{C}$ given in (20). Clearly (24) implies that $\mathbf{dist}(f) \leqslant 1+\varepsilon$, as required.

To prove the right hand inequality in (24), suppose that $i,j \in \{1,\ldots,n\}$ satisfy $i<j$ and consider the coupling $\pi \in \Pi(f(x_i),f(x_j))$ given by

$$\pi \stackrel{\mathrm{def}}{=} \frac{1}{N}\bigg(\sum_{t=1}^{5}\sum_{s=0}^{K-1}\delta_{\left(Q_s^t(i,j),Q_{s+1}^t(i,j)\right)} + \delta_{\left(Q_K^2(i,j),Q_0^3(i,j)\right)} + \sum_{u\in\mathcal{C}\smallsetminus\mathcal{B}_{ij}}\delta_{(u,u)}\bigg), \tag{25}$$

where for (25) recall (19) and (20). The meaning of (25) is simple: the supports of $f(x_i)$ and $f(x_j)$ equal $\mathcal{C}_i$ and $\mathcal{C}_j$, respectively, where we recall (21). Note that $\mathcal{C}_i \smallsetminus \mathcal{C}_j = \{Q_0^1(i,j)\}$ and $\mathcal{C}_j \smallsetminus \mathcal{C}_i = \{Q_K^5(i,j)\}$, where we recall (13) and (17). So, the coupling π in (25) corresponds to shifting the points in $\mathcal{B}_{ij}$ from the support of $f(x_i)$ to the support of $f(x_j)$ while keeping the points in $\mathcal{C}\smallsetminus\mathcal{B}_{ij}$ unchanged.

Now, recalling the definitions (13), (14), (15), (16) and (17),

$$\begin{aligned}\mathsf{W}_p(f(x_i),f(x_j))^p &\leqslant \iint_{\mathbb{R}^3\times\mathbb{R}^3}\|x-y\|_2^p\,\mathrm{d}\pi(x,y)\\ &= \frac{1}{N}\sum_{t=1}^{5}\sum_{s=0}^{K-1}\left\|Q_s^t(i,j)-Q_{s+1}^t(i,j)\right\|_2^p + \frac{\|Q_K^2(i,j)-Q_0^3(i,j)\|_2^p}{N}.\end{aligned} \tag{26}$$

Note that if $s \in \{0,\ldots,K-1\}$ then by (13), (14), (16), (17) we have

$$\begin{aligned} t\in\{1,5\} &\implies \left\|Q_s^t(i,j)-Q_{s+1}^t(i,j)\right\|_2 = \frac{M\phi(i,j)}{mK} \leqslant \frac{Mn^2}{mK},\\ t\in\{2,4\} &\implies \left\|Q_s^t(i,j)-Q_{s+1}^t(i,j)\right\|_2 = \frac{M}{mK}.\end{aligned} \tag{27}$$

Also, by (14) and (15) we have

$$\left\|Q_K^2(i,j)-Q_0^3(i,j)\right\|_2 = \frac{d_X(x_i,x_j)^{\frac{1}{p}}}{m}. \tag{28}$$

Finally, by (15) for every $s \in \{0,\ldots,K-1\}$ we have

$$\left\|Q_s^3(i,j)-Q_{s+1}^3(i,j)\right\|_2 = \frac{M(j-i)}{mK} - \frac{d_X(x_i,x_j)^{\frac{1}{p}}}{mK} \leqslant \frac{Mn}{mK}, \tag{29}$$

where in the first step of (29) we used the fact that $M(j-i)-d_X(x_i,x_j)^{1/p} \geqslant 0$, which holds true by the definition of M in (12) because $j-i\geqslant 1$. A substitution of (27), (28) and (29) into (26) yields the estimate

$$\mathsf{W}_p(f(x_i),f(x_j))^p \leqslant \frac{d_X(x_i,x_j)}{m^pN}+\frac{5K}{N}\left(\frac{Mn^2}{mK}\right)^p$$
$$=\left(1+\frac{5M^pn^{2p}}{K^{p-1}d_X(x_i,x_j)}\right)\frac{d_X(x_i,x_j)}{m^pN}\leqslant (1+p\varepsilon)\frac{d_X(x_i,x_j)}{m^pN},$$

where we used the fact that by the definition of m in (12) we have $m^p \leqslant d_X(x_i,x_j)$, and the lower bound on K that is assumed in (23). This implies the right hand inequality in (24) because $1+p\varepsilon \leqslant (1+\varepsilon)^p$.

Passing now to the proof of the left hand inequality in (24), we need to prove that for every $i,j\in\{1,\ldots,n\}$ with $i<j$ we have

$$\forall\, \pi\in \Pi(f(x_i),f(x_j)),\qquad \iint_{\mathbb{R}^3\times\mathbb{R}^3}\|x-y\|_2^p \mathrm{d}\pi(x,y)\geqslant \frac{d_X(x_i,x_j)}{m^pN}. \tag{30}$$

Note that we still did not use the triangle inequality for d_X, but this will be used in the proof of (30). Also, the reason why we are dealing with $\mathcal{P}_p(\mathbb{R}^3)$ rather than $\mathcal{P}_p(\mathbb{R}^2)$ will become clear in the ensuing argument.

Recall that the measures $f(x_i)$ and $f(x_j)$ are uniformly distributed over sets of the same size, and their supports $\mathcal{C}_i$ and $\mathcal{C}_j$ (respectively) satisfy $\mathcal{C}_i\triangle\mathcal{C}_j=\{(Mi/m)e_1,(Mj/m)e_1\}$. Since the set of all doubly stochastic matrices is the convex hull of the permutation matrices, and every permutation is a product of disjoint cycles, it follows that it suffices to establish the validity of (30) when $\pi=\frac{1}{N}\sum_{\ell=1}^L\delta_{(u_{\ell-1},u_\ell)}$ for some $L\in\{1,\ldots,N\}$ and $u_1,\ldots u_{L-1}\in\mathcal{C}$, where we set $u_0=(Mi/m)e_1$ and $u_L=(Mj/m)e_1$. With this notation, our goal is to show that

$$\frac{1}{N}\sum_{\ell=1}^L\|u_\ell-u_{\ell-1}\|_2^p\geqslant \frac{d_X(x_i,x_j)}{m^pN}. \tag{31}$$

For every $a\in\{1,\ldots,n\}$ define $\mathcal{S}_a\subseteq\mathbb{R}^3$ by $\mathcal{S}_a\stackrel{\mathrm{def}}{=}\mathcal{S}_a^1\cup\mathcal{S}_a^2$, where

$$\mathcal{S}_a^1\stackrel{\mathrm{def}}{=}\bigcup_{b=a+1}^n\bigcup_{s=0}^K\left\{Q_s^1(a,b),Q_s^2(a,b)\right\}, \tag{32}$$

and

$$\mathcal{S}_a^2\stackrel{\mathrm{def}}{=}\bigcup_{c=1}^{a-1}\bigcup_{s=0}^K\left\{Q_s^3(c,a),Q_s^4(c,a),Q_s^5(c,a)\right\}. \tag{33}$$

Thus, recalling (18), the sets $\mathcal{S}_1,\ldots,\mathcal{S}_n$ form a partition of $\mathcal{B}$ and $a\in\mathcal{S}_a$ for every $a\in\{1,\ldots,n\}$. For every $\ell\in\{0,\ldots,L\}$ let $a(\ell)$ be the unique element of $\{1,\ldots,n\}$ for which $u_\ell\in\mathcal{S}_{a(\ell)}$. Then $a(0)=i$ and $a(L)=j$. The left hand side of (31) can be bounded from below as follows

$$\frac{1}{N}\sum_{\ell=1}^L\|u_\ell-u_{\ell-1}\|_2^p\geqslant \frac{1}{N}\sum_{\ell=1}^L\min_{\substack{u\in\mathcal{S}_{a(\ell-1)}\\ v\in\mathcal{S}_{a(\ell)}}}\|u-v\|_2^p. \tag{34}$$

We shall show that

$$\forall\, a,b\in\{1,\ldots,n\},\ \forall (u,v)\in\mathcal{S}_a\times\mathcal{S}_b,\qquad \|u-v\|_2^p\geqslant\frac{d_X(x_a,x_b)}{m^p}. \tag{35}$$

The validity of (35) implies the required estimate (31) because, by (34), it follows from (35) and the triangle inequality for d_X that

$$\frac{1}{N}\sum_{\ell=1}^L\|u_\ell-u_{\ell-1}\|_2^p\geqslant\frac{1}{N}\sum_{\ell=1}^L\frac{d_X\big(x_{a(\ell-1)},x_{a(\ell)}\big)}{m^p}\geqslant\frac{d_X(x_i,x_j)}{m^pN}.$$

It remains to justify (35). Suppose that $a,b\in\{1,\ldots,n\}$ satisfy $a<b$ and $(u,v)\in\mathcal{S}_a\times\mathcal{S}_b$. Write $u=Q_s^t(c,d)$ and $v=Q_\sigma^\tau(\gamma,\delta)$ for some $s,\sigma\in\{0,\ldots,K\}$, $t,\tau\in\{1,\ldots,5\}$ and $c,d,\gamma,\delta\in\{1,\ldots,n\}$.

We shall check below, via a direct case analysis, that the absolute value of one of the three coordinates of $u-v$ is either at least M/m or at least $d_X(x_a,x_b)^{1/p}/m$. Since by the definition of M in (12) we have $M\geqslant d_X(x_a,x_b)^{1/p}$, this assertion will imply (35).

Suppose first that $t,\tau\in\{1,2,4,5\}$. By comparing (32), (33) with (13), (14), (16), (17) we see that $\langle u,e_1\rangle = Ma/m$ and $\langle v,e_1\rangle=Mb/m$. Since $b-a\geqslant 1$, this implies that $\langle u-v,e_1\rangle\geqslant M/m$, as required.

If $t=\tau=3$ then by (33) we necessarily have $d=a$ and $\delta=b$. Hence $(c,d)\neq(\gamma,\delta)$ and therefore $|\phi(c,d)-\phi(\gamma,\delta)|\geqslant 1$, since ϕ is a bijection between $\{1,\ldots,n\}\times\{1,\ldots,n\}$ and $\{1,\ldots,n^2\}$. By (15) we therefore have $|\langle u-v,e_2\rangle|\geqslant M/m$, as required.

It remains to treat the case $t\neq\tau$ and $3\in\{t,\tau\}$. If $\{t,\tau\}\subseteq\{1,3,5\}$ then by contrasting (15) with (13) and (17) we see that the third coordinate of one of the vectors u,v vanishes while the third coordinate of the other vector equals M/m. Therefore $|\langle u-v,e_3\rangle|\geqslant M/m$, as required. The only remaining case is $\{t,\tau\}\subseteq\{2,3,4\}$. In this case $|\langle u-v,e_2\rangle| = M|\phi(c,d)-\phi(\gamma,\delta)|/m$, by (15), (14), (16). So, if $(c,d)\neq(\gamma,\delta)$ then $|\phi(c,d)-\phi(\gamma,\delta)|\geqslant 1$, and we are done. We may therefore assume that $c=\gamma$ and $d=\delta$. Observe that by (33) if $\{t,\tau\}=\{3,4\}$ then $\{d,\delta\}=\{a,b\}$, which contradicts $d=\delta$. So, we also necessarily have $\{t,\tau\}=\{2,3\}$, in which case, since $a<b$, by (32) and (33) we see that $c=\gamma=a$ and $d=\delta=b$. By interchanging the labels s and σ if necessary, we may assume that $u=Q_\sigma^2(a,b)$ and $v=Q_s^3(a,b)$. By (14) and (15) we therefore have

$$\langle v-u,e_1\rangle = \frac{M(s(b-a)+Ka)}{mK}+\frac{(K-s)d_X(x_a,x_b)^{\frac1p}}{mK}-\frac{Ma}{m} = \frac{d_X(x_a,x_b)^{\frac1p}}{m}+\frac{sM(b-a)-sd_X(x_a,x_b)^{\frac1p}}{mK}\geqslant \frac{d_X(x_a,x_b)^{\frac1p}}{m},$$

where we used the fact that by (12) we have $M\geqslant d_X(x_a,x_b)^{1/p}$, and that $b-a\geqslant 1$. This concludes the verification of the remaining case of (35), and hence the proof of Lemma 15 is complete. □

3. Sharpness of Theorem 1

The results of this section rely crucially on K. Ball's notion [5] of Markov type. We shall start by briefly recalling the relevant background on this important invariant of metric spaces, including variants and notation from [67] that will be used below. Let $\{Z_t\}_{t=0}^\infty$ be a Markov chain on the state space $\{1,\ldots,n\}$ with transition probabilities $a_{ij}=\Pr[Z_{t+1}=j|Z_t=i]$ for every $i,j\in\{1,\ldots,n\}$. $\{Z_t\}_{t=0}^\infty$ is said to be stationary if $\pi_i=\Pr[Z_t=i]$ does not depend on $t\in\{1,\ldots,n\}$ and it is said to be reversible if $\pi_i a_{ij}=\pi_j a_{ji}$ for every $i,j\in\{1,\ldots,n\}$.

Let $\{Z_t'\}_{t=0}^\infty$ be the Markov chain that starts at Z_0 and then evolves independently of $\{Z_t\}_{t=0}^\infty$ with the same transition probabilities. Thus $Z_0'=Z_0$ and conditioned on Z_0 the random variables Z_t and Z_t' are independent and identically distributed. We note for future use that if $\{Z_t\}_{t=0}^\infty$ as above is stationary and reversible then for every symmetric function $\psi:\{1,\ldots,n\}\times\{1,\ldots,n\}\to\mathbb{R}$ and every $t\in\mathbb{N}$ we have

$$\mathbb{E}\big[\psi(Z_t,Z_t')\big]=\mathbb{E}\big[\psi(Z_{2t},Z_0)\big]. \tag{36}$$

This is a consequence of the observation that, by stationarity and reversibility, conditioned on the random variable Z_t the random variables Z_0 and Z_{2t} are independent and identically distributed. Denoting $A=(a_{ij})\in M_n(\mathbb{R})$, the validity of (36) can be alternatively checked directly as follows.

$$\mathbb{E}\left[\psi(Z_t,Z_t')\right]=\mathbb{E}\big[\mathbb{E}\left[\psi(Z_t,Z_t')|Z_0\right]\big]=\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\pi_iA_{ij}^tA_{ik}^t\psi(j,k)$$
$$\stackrel{(\star)}{=}\sum_{j=1}^n\sum_{k=1}^n\pi_j\bigg(\sum_{i=1}^nA_{ji}^tA_{ik}^t\bigg)\psi(j,k)=\sum_{j=1}^n\sum_{k=1}^n\pi_jA_{jk}^{2t}\psi(j,k),\quad(37)$$

where $(\star)$ uses the reversibility of the Markov chain $\{Z_t\}_{t=0}^\infty$ through the validity of $\pi_iA_{ij}^t=\pi_jA_{ji}^t$ for every $i,j\in\{1,\ldots,n\}$. The final term in (37) equals the right hand side of (36), as required.

Given $p\in[1,\infty)$, a metric space (X,d_X) and $m\in\mathbb{N}$, the Markov type p constant of (X,d_X) at time m, denoted $M_p(X,d_X;m)$ (or simply $M_p(X;m)$ if the metric is clear from the context) is defined to be the infimum over those $M\in(0,\infty)$ such that for every $n\in\mathbb{N}$, every stationary reversible Markov chain $\{Z_t\}_{t=0}^\infty$ with state space $\{1,\ldots,n\}$, and every $f:\{1,\ldots,n\}\to X$ we have

$$\mathbb{E}\big[d_X(f(Z_m),f(Z_0))^p\big]\leqslant M^pm\mathbb{E}\big[d_X(f(Z_1),f(Z_0))^p\big].$$

Observe that by the triangle inequality we always have

$$M_p(X;m)\leqslant m^{1-\frac1p}.$$

As we shall explain below, any estimate of the form $M_p(X;m)\lesssim_X m^\theta$ for $\theta<1-1/p$ is a nontrivial obstruction to the embeddability of certain metric spaces into X, but it is especially important (e.g. for Lipschitz extension theory [5]) to single out the case when $M_p(X;m)\lesssim_X 1$. Specifically, (X,d_X) is said to have Markov type p if

$$M_p(X,d_X)\stackrel{\mathrm{def}}{=}\sup_{m\in\mathbb{N}}M_p(X,d_X;m)<\infty.$$

$M_p(X,d_X)$ is called the Markov type p constant of (X,d_X), and it is often denoted simply $M_p(X)$ if the metric is clear from the context.

The Markov type of many important classes of metric spaces is satisfactorily understood, though some fundamental questions remain open; see Section 4 of the survey [66] and the references therein, as well as more recent progress in e.g. [22]. Here we study this notion in the context of Wasserstein spaces. The link of Markov type to the nonembeddability of snowflakes is simple, originating in an idea of [51]. This is the content of the following lemma.

Lemma 16. *Fix a metric space (Y,d_Y), $m\in\mathbb{N}$, $K,p\in[1,\infty)$ and $\theta\in[0,1]$. Suppose that*

$$M_p(Y;m)\leqslant Km^{\frac{\theta(p-1)}{p}}.\quad(38)$$

Denote $n=2^{4m}$. Then there exists an n-point metric space (X,d_X) such that

$$\alpha\in\left[\frac{1+\theta(p-1)}{p},1\right]\implies c_Y(X,d_X^\alpha)\gtrsim\frac1K(\log n)^{\alpha-\frac{1+\theta(p-1)}{p}}.$$

Proof. Take $(X,d_X)=(\{0,1\}^{4m},\|\cdot\|_1)$, i.e., X is the $4m$-dimensional discrete hypercube, equipped with the Hamming metric. Thus $|X|=n$. Let $\{Z_t\}_{t=0}^\infty$ be the standard random walk on X, with Z_0 distributed uniformly over X. Suppose that $f:X\to Y$ satisfies

$$\forall\,x,y\in X,\qquad s\|x-y\|_1^\alpha\leqslant d_Y(f(x),f(y))\leqslant Ds\|x-y\|_1^\alpha\quad(39)$$

for some $s,D\in(0,\infty)$. Our goal is to bound D from below. By the definition of $M_p(Y;m)$,

$$\mathbb{E}\big[d_Y(f(Z_m),f(Z_0))^p\big]\stackrel{(38)}{\leqslant}K^pm^{1+\theta(p-1)}\mathbb{E}\big[d_Y(f(Z_1),f(Z_0))^p\big].\quad(40)$$

By the right hand inequality in (39) we have

$$\mathbb{E}\big[d_Y(f(Z_1),f(Z_0))^p\big]\leqslant D^ps^p\mathbb{E}\big[\|Z_1-Z_0\|_1^{\alpha p}\big]=D^ps^p.\quad(41)$$

At the same time, it is simple to see (and explained explicitly in e.g. [71] or [66, Section 9.4]) that $\mathbb{E}\big[\|Z_m - Z_0\|_1^{\alpha p}\big] \geqslant (\eta m)^{\alpha p}$ for some universal constant $\eta \in (0,1)$. Hence,

$$\mathbb{E}\big[d_Y(f(Z_m), f(Z_0))^p\big] \overset{(39)}{\geqslant} s^p \mathbb{E}\big[\|Z_m - Z_0\|_1^{\alpha p}\big] \gtrsim s^p (\eta m)^{\alpha p}. \tag{42}$$

The only way for (41) and (42) to be compatible with (40) is if

$$D \gtrsim \frac{1}{K} m^{\alpha - \frac{1+\theta(p-1)}{p}} \asymp \frac{1}{K} (\log n)^{\alpha - \frac{1+\theta(p-1)}{p}}. \qquad \square$$

Remark 17. In Lemma 16 we took the metric space X to be a discrete hypercube, but similar conclusions apply to snowflakes of expander graphs and graphs with large girth [51], as well as their subsets [7] and certain discrete groups [4, 68, 69] (see also [66, Section 9.4]). We shall not attempt to state here the wider implications of the assumption (38) to the nonembeddability of snowflakes, since the various additional conclusions follow mutatis mutandis from the same argument as above, and Lemma 16 as currently stated suffices for the proof of Theorem 2.

Remark 18. Since the proof of Lemma 16 applied the Markov type p assumption (38) to the discrete hypercube, it would have sufficed to work here with a classical weaker bi-Lipschitz invariant due to Enflo [24], called Enflo type. Such an obstruction played a role in ruling out certain snowflake embeddings in [26] (in a different context), though the fact that the argument of [26] could be cast in the context of Enflo type was proved only later [76, Proposition 5.3]. Here we work with Markov type rather than Enflo type because the proof below for Wasserstein spaces yields this stronger conclusion without any additional effort.

The following lemma is a variant of [76, Lemma 4.1].

Lemma 19. *Fix $p \in [1,\infty)$ and $\theta \in [1/p, 1]$. Suppose that (X, d_X) is a metric space such that for every two X-valued independent and identically distributed finitely supported random variables Z, Z' and every $x \in X$ we have*

$$\mathbb{E}\big[d_X(Z, Z')^p\big] \leqslant 2^{\theta p} \mathbb{E}\big[d_X(Z, x)^p\big]. \tag{43}$$

Then for every $k \in \mathbb{N}$ we have

$$M_p(X; 2^k) \leqslant 2^{k\left(\theta - \frac{1}{p}\right)}. \tag{44}$$

Proof. Fix $n \in \mathbb{N}$, a stationary reversible Markov chain $\{Z_t\}_{t=0}^{\infty}$ with state space $\{1,\ldots,n\}$, and $f : \{1,\ldots,n\} \to X$. Recalling (36) with $\psi(i,j) = d_X(f(i), f(j))^p$, for every $t \in \mathbb{N}$ we have

$$\begin{aligned}\mathbb{E}\big[d_X(Z_{2t}, Z_0)^p\big] \overset{(36)}{=} \mathbb{E}\big[d_X(Z_t, Z_t')^p\big] &\overset{(43)}{\leqslant} 2^{\theta p} \mathbb{E}\big[d_X(Z_t, Z_0)^p\big] \\ &\leqslant 2^{\theta p - 1} M_p(X; t)^p \cdot 2t \mathbb{E}\big[d_X(Z_1, Z_0)^p\big], \end{aligned} \tag{45}$$

where the last step of (45) uses the definition of $M_p(X; t)$. By the definition of $M_p(X; 2t)$, we have thus proved that

$$M_p(X; 2t) \leqslant 2^{\theta - \frac{1}{p}} M_p(X; t),$$

so (44) follows by induction on k. $\square$

Corollary 20 below follows from Lemma 16 and Lemma 19. Specifically, under the assumptions and notation of Lemma 19, use Lemma 16 with m replaced by 2^k and θ replaced by $(\theta p - 1)/(p-1)$.

Corollary 20. *Fix $p \in [1,\infty)$ and $\theta \in [1/p, 1]$. Suppose that (X, d_X) is a metric space that satisfies the assumptions of Lemma 19. Then for arbitrarily large $n \in \mathbb{N}$ there exists an n-point metric space (Y, d_Y) such that for every $\alpha \in [\theta, 1]$ we have*

$$c_X\left(Y, d_Y^{\alpha}\right) \gtrsim (\log n)^{\alpha - \theta}.$$

The link between the above discussion and embeddings of snowflakes of metrics into Wasserstein spaces is explained in the following lemma, which is a variant of [89, Proposition 2.10].

Lemma 21. *Fix $p \in [1,\infty)$ and $\theta \in [1/p,1]$. Suppose that (X,d_X) is a metric space that satisfies the assumptions of Lemma 19, i.e., inequality (43) holds true for X-valued random variables. Then the same inequality holds true in the metric space $(\mathcal{P}_p(X),\mathsf{W}_p)$ as well, i.e., for every two $\mathcal{P}_p(X)$-valued and identically distributed finitely supported random variables $\mathfrak{M},\mathfrak{M}'$ and every $\mu \in \mathcal{P}_p(X)$,*

$$\mathbb{E}\big[\mathsf{W}_p(\mathfrak{M},\mathfrak{M}')^p\big] \leqslant 2^{\theta p}\mathbb{E}\big[\mathsf{W}_p(\mathfrak{M},\mu)^p\big].$$

Proof. Suppose that the distribution of $\mathfrak{M}$ equals $\sum_{i=1}^n q_i\delta_{\mu_i}$ for some $\mu_1,\ldots,\mu_n \in \mathcal{P}_p(X)$ and $q_1,\ldots,q_n \in [0,1]$ with $\sum_{i=1}^n q_i = 1$. Our goal is to show that

$$\sum_{i=1}^n\sum_{j=1}^n q_iq_j\mathsf{W}_p(\mu_i,\mu_j)^p \leqslant 2^{\theta p}\sum_{i=1}^n q_i\mathsf{W}_p(\mu_i,\mu)^p. \tag{46}$$

The finitely supported probability measures are dense in $(\mathcal{P}_p(X),\mathsf{W}_p)$ (see [83, 94]), so it suffices to prove (46) when there exists $N\in\mathbb{N}$ and points $x_{ik},x_k \in X$ for every $(i,k)\in\{1,\ldots,n\}\times\{1,\ldots,N\}$ such that we have $\mu = \frac{1}{N}\sum_{k=1}^N \delta_{x_k}$ and $\mu_i = \frac{1}{N}\sum_{k=1}^N\delta_{x_{ik}}$ for every $i\in\{1,\ldots,n\}$. Let $\{\sigma_i\}_{i=1}^n \subseteq S_N$ be permutations of $\{1,\ldots,N\}$ that induce optimal couplings of the pairs (μ,μ_i), i.e.,

$$\forall\, i\in\{1,\ldots,n\},\qquad \mathsf{W}_p(\mu_i,\mu)^p = \frac{1}{N}\sum_{k=1}^N d_X(x_{i\sigma_i(k)},x_k)^p. \tag{47}$$

Since the measure $\frac{1}{N}\sum_{k=1}^N \delta_{(x_{i\sigma_i(k)},x_{j\sigma_j(k)})}$ is a coupling of (μ_i,μ_j),

$$\forall\, i,j\in\{1,\ldots,n\},\qquad \mathsf{W}_p(\mu_i,\mu_j)^p \leqslant \frac{1}{N}\sum_{k=1}^N d_X(x_{i\sigma_i(k)},x_{j\sigma_j(k)})^p. \tag{48}$$

Consequently,

$$\begin{aligned}\sum_{i=1}^n\sum_{j=1}^n q_iq_j\mathsf{W}_p(\mu_i,\mu_j)^p &\overset{(48)}{\leqslant} \frac{1}{N}\sum_{k=1}^N\sum_{i=1}^n\sum_{j=1}^n q_iq_jd_X(x_{i\sigma_i(k)},x_{j\sigma_j(k)})^p\\ &\overset{(43)}{\leqslant} \frac{2^{\theta p}}{N}\sum_{k=1}^N\sum_{i=1}^n\sum_{j=1}^n q_iq_jd_X(x_{i\sigma_i(k)},x_k)^p \overset{(47)}{=} 2^{\theta p}\sum_{i=1}^n q_i\mathsf{W}_p(\mu_i,\mu)^p. \qquad\square\end{aligned}$$

Proof of Theorem 2. Let (Ω,μ) be a probability space. For $p\in[1,\infty]$ define $T: L_p(\mu)\to L_p(\mu\times\mu)$ by $Tf(x,y)=f(x)-f(y)$. Then clearly $\|T\|_{L_p(\mu)\to L_p(\mu\times\mu)}\leqslant 2$ for $p\in\{1,\infty\}$ and

$$\forall\, f\in L_2(\mu),\qquad \|Tf\|_{L_2(\mu\times\mu)}^2 = 2\|f\|_{L_2(\mu)}^2 - 2\Big(\int_\Omega f\,\mathrm{d}\mu\Big)^2 \leqslant 2\|f\|_{L_2(\mu)}^2.$$

Or $\|T\|_{L_2(\mu)\to L_2(\mu\times\mu)}\leqslant\sqrt{2}$. So, by the Riesz–Thorin theorem (e.g. [27]),

$$p\in[1,2] \implies \|T\|_{L_p(\mu)\to L_p(\mu\times\mu)} \leqslant 2^{\frac{1}{p}}, \tag{49}$$

and

$$p\in[2,\infty] \implies \|T\|_{L_p(\mu)\to L_p(\mu\times\mu)} \leqslant 2^{1-\frac{1}{p}}. \tag{50}$$

Switching to probabilistic terminology, the estimates (49) and (50) say that if Z,Z' are i.i.d. random variables then $\mathbb{E}\big[|Z-Z'|^p\big]\leqslant 2\mathbb{E}\big[|Z|^p\big]$ when $p\in[1,2]$ and $\mathbb{E}\big[|Z-Z'|^p\big]\leqslant 2^{p-1}\mathbb{E}\big[|Z|^p\big]$ when

$p\in[2,\infty)$. By applying this to the random variables $Z-a, Z'-a$ for every $a\in\mathbb{R}$, we deduce that the real line (with its usual metric) satisfies (43) with

$$\theta=\theta_p \stackrel{\mathrm{def}}{=} \max\left\{\frac{1}{p},1-\frac{1}{p}\right\}. \tag{51}$$

Invoking this statement coordinate-wise shows that $\ell_p^3=(\mathbb{R}^3,\|\cdot\|_p)$ satisfies (43) with $\theta=\theta_p$. Lemma 21 therefore implies that $(\mathcal{P}_p(\ell_p^3),\mathsf{W}_p)$ also satisfies (43) with $\theta=\theta_p$. Hence, by Corollary 20 for arbitrarily large $n\in\mathbb{N}$ there exists an n-point metric space (Y,d_Y) such that for every $\alpha\in(\theta_p,1]$,

$$c_{(\mathcal{P}_p(\ell_p^3),\mathsf{W}_p)}(Y,d_Y^\alpha)\gtrsim(\log n)^{\alpha-\theta_p}=\begin{cases}(\log n)^{\alpha-\frac{1}{p}} & \text{if } p\in(1,2],\\ (\log n)^{\alpha+\frac{1}{p}-1} & \text{if } p\in(2,\infty).\end{cases}$$

Since the ℓ_p norm on $\mathbb{R}^3$ is $\sqrt{3}$-equivalent to the ℓ_2 norm on $\mathbb{R}^3$,

$$c_{(\mathcal{P}_p(\ell_p^3),\mathsf{W}_p)}(Y,d_Y^\alpha)\asymp c_{(\mathcal{P}_p(\ell_2^3),\mathsf{W}_p)}(Y,d_Y^\alpha),$$

thus completing the proof of Theorem 2. □

Remark 22. In the proof of Theorem 2 we chose to check the validity of (43) with $\theta=\theta_p$ given in (51) using an interpolation argument since it is very short. But, there are different proofs of this fact: when $p\in[1,2)$ one could start from the trivial case $p=2$, and then pass to general $p\in[1,2)$ by invoking the classical fact [86] that the metric space $(\mathbb{R},|x-y|^{p/2})$ admits an isometric embedding into Hilbert space. Alternatively, in [64, Lemma 3] this is proved via a direct computation.

Question 23. As discussed in the Introduction, it seems plausible that Theorem 1 and Theorem 2 are not sharp when $p\in(2,\infty)$. Specifically, we conjecture that there exist $D_p\in[1,\infty)$ such that for every finite metric space (X,d_X) we have

$$c_{\mathcal{P}_p(\mathbb{R}^3)}\Big(X,\sqrt{d_X}\Big)\leqslant D_p. \tag{52}$$

Perhaps (52) even holds true with $D_p=1$. As discussed in Remark 11, since L_2 admits an isometric embedding into L_p (see e.g. [96]), the perceived analogy between Wasserstein p spaces and L_p spaces makes it natural to ask whether or not $(\mathcal{P}_2(\mathbb{R}^3),\mathsf{W}_2)$ admits a bi-Lipschitz embedding into $(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)$. If the answer to this question were positive then (52) would hold true by virtue of the case $p=2$ of Theorem 1. We also conjecture that the lower bound of Theorem 2 could be improved when $p>2$ to state that for arbitrarily large $n\in\mathbb{N}$ there exists an n-point metric space (Y,d_Y) such that for every $\alpha\in(1/2,1]$,

$$c_{(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)}(Y,d_Y^\alpha)\gtrsim_p(\log n)^{\alpha-\frac{1}{2}}. \tag{53}$$

It was shown in [70] that L_p has Markov type 2 for every $p\in(2,\infty)$. We therefore ask whether or not $(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)$ has Markov type 2 for every $p\in(2,\infty)$. A positive answer to this question would imply that the lower bound (53) is indeed achievable. For this purpose it would also suffice to show that for every $p\in(2,\infty)$ and $k\in\mathbb{N}$ we have

$$M_p((\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p);2^k)\lesssim_p 2^{k\left(\frac{1}{2}-\frac{1}{p}\right)}. \tag{54}$$

Proving (54) may be easier than proving that $M_2(\mathcal{P}_p(\mathbb{R}^3),\mathsf{W}_p)<\infty$, since the former involves arguing about the pth powers of Wasserstein p distances while the latter involves arguing about Wasserstein p distances squared. Note that $M_p(L_p;m)\lesssim\sqrt{p}m^{1/2-1/p}$ by [70] (see also [67, Theorem 4.3]), so the L_p-version of (54) is indeed valid.

We end this section by showing how Lemma 19 implies bounds on the Markov type p constant $M_p(X;t)$ for any time $t\in\mathbb{N}$, and not only when $t=2^k$ for some $k\in\mathbb{N}$ as in (44). For the purpose of proving Theorem 1, Lemma 19 suffices as stated, so the ensuing discussion is included for completeness, and could be skipped by those who are interested only in the proof of Theorem 2.

The case $p = 2$ and $\theta = 1/2$ of Lemma 19 corresponds to proving that metric spaces that are nonnegatively curved in the sense of Alexandrov have Markov type 2: this was established by Ohta in [76], whose work inspired the arguments that were presented above. Specifically, Ohta showed in [76] how to pass from (44) with $p = 2$ and $\theta = 1/2$ (i.e., $M_2(X, 2^k) \leqslant 1$ for every $k \in \mathbb{N}$) to $M_2(X) \leqslant \sqrt{6} = 2.449...$, and he also included in [76] an argument of Naor and Peres that improves this to $M_2(X) \leqslant 1 + \sqrt{2} = 2.414...$ Below we further refine the latter argument, yielding the best known estimate on the Markov type 2 constant of Alexandrov spaces of nonnegative curvature; see (57) below. This constant is of interest since it was shown in [77] that if (X, d_X) is a geodesic metric space with $M_2(X) = 1$ then X is nonnegatively curved in the sense of Alexandrov. It is plausible that, conversely, $M_2(X) = 1$ if X is nonnegatively curved in the sense of Alexandrov, but, as noted in [77], this seems to be unknown even for the circle $X = S^1$.

For every $\theta \in (0, 1]$ define $\phi_\theta : [0, 1] \to \mathbb{R}$ by

$$\forall\, s \in [0, 1], \qquad \phi_\theta(s) \stackrel{\mathrm{def}}{=} s^\theta - (1 - s)^\theta. \tag{55}$$

Then $\phi_\theta([0, 1]) = [-1, 1]$ and since $\phi'_\theta(s) = \theta s^{\theta-1} + \theta(1 - s)^{\theta-1} > 0$, the inverse ϕ_θ^{-1} is well-defined and increasing on $[-1, 1]$. The following elementary numerical lemma will be used later.

Lemma 24. *For all $\theta \in (0, 1]$ there is a unique $c(\theta) \in (1, \infty)$ satisfying*

$$c(\theta) = \frac{c(\theta)\phi_\theta^{-1}\left(\frac{2^\theta-1}{c(\theta)}\right)^\theta + 1}{\left(\phi_\theta^{-1}\left(\frac{2^\theta-1}{c(\theta)}\right) + 1\right)^\theta}. \tag{56}$$

Proof. The identity (56) is equivalent to $h_\theta(c(\theta)) = 1$, where for every $s > 0$ and $c \in [1, \infty)$ we set

$$\psi_\theta(s) \stackrel{\mathrm{def}}{=} (s + 1)^\theta - s^\theta \qquad \text{and} \qquad h_\theta(c) \stackrel{\mathrm{def}}{=} c\psi_\theta\left(\phi_\theta^{-1}\left(\frac{2^\theta - 1}{c}\right)\right).$$

Observe that because $\theta \in (0, 1]$ we have $\psi_\theta(s) < 1$ for every $s > 0$. Hence $h_\theta(c) < c$ for every $c \in (0, \infty)$, and in particular $h_\theta(1) < 1$. Moreover, $\phi_\theta^{-1}(0) = 1/2$, so that

$$\lim_{c\to\infty} \psi_\theta\left(\phi_\theta^{-1}\left(\frac{2^\theta - 1}{c}\right)\right) = \psi_\theta\left(\frac{1}{2}\right) = \frac{3^\theta}{2^\theta} - \frac{1}{2^\theta} > 0.$$

Hence $\lim_{c\to\infty} h_\theta(c) = \infty$. It follows by continuity that there exists $c \in (0, \infty)$ such that $h_\theta(c) = 1$. To prove the uniqueness of such $c > 1$, it suffice to show that h_θ is increasing on $(0, \infty)$. Now,

$$h'_\theta(c) = \psi_\theta\left(\phi_\theta^{-1}\left(\frac{2^\theta - 1}{c}\right)\right) - \frac{2^\theta - 1}{c} \cdot \frac{\psi'_\theta\left(\phi_\theta^{-1}\left(\frac{2^\theta-1}{c}\right)\right)}{\phi'_\theta\left(\phi_\theta^{-1}\left(\frac{2^\theta-1}{c}\right)\right)} = \frac{\phi'_\theta(y)\psi_\theta(y) - \phi_\theta(y)\psi'_\theta(y)}{\phi'_\theta(y)},$$

where we write $y = \phi_\theta^{-1}((2^\theta - 1)/c)$. Since ϕ_θ is increasing, it therefore suffices to show that $\phi'_\theta(y)\psi_\theta(y) - \phi_\theta(y)\psi'_\theta(y) > 0$ for all $y \in (0, 1)$. One directly computes that

$$\phi'_\theta(y)\psi_\theta(y) - \phi_\theta(y)\psi'_\theta(y) = \theta \cdot \frac{2y^{1-\theta} + (1 - y)^{1-\theta} - (1 + y)^{1-\theta}}{y^{1-\theta}(1 - y^2)^{1-\theta}}.$$

It remains to note that by the subadditivity of $t \mapsto t^{1-\theta}$ we have

$$(1 + y)^{1-\theta} \leqslant (1 - y)^{1-\theta} + (2y)^{1-\theta} \leqslant (1 - y)^{1-\theta} + 2y^{1-\theta}.$$

□

Lemma 25. *Fix $p \in [1,\infty)$ and $\theta \in [1/p,1]$. Suppose that (X,d_X) is a metric space that satisfies the assumptions of Lemma 19, i.e., inequality (43) holds true for X-valued random variables. Then*

$$\forall\, t \in \mathbb{N}, \qquad M_p(X;t) \leqslant c(\theta) t^{\theta-\frac{1}{p}},$$

where $c(\theta)$ is from Lemma 24. Thus, if $\theta = 1/p$ then X has Markov type p with $M_p(X) \leqslant c(1/p)$.

Because, by the Lang–Schroeder–Sturm inequality (6), Alexandrov spaces of nonnegative curvature satisfy the assumption of Lemma 19 with $p = 2$ and $\theta = 1/2$, we have the following corollary. Note that $c(1/2)$ can be computed explicitly by solving the equation (56).

Corollary 26. *Suppose that (X,d_X) is nonnegatively curved in the sense of Alexandrov. Then the Markov type 2 constant of X satisfies*

$$M_2(X) \leqslant c\left(\frac{1}{2}\right) = \sqrt{1+\sqrt{2}+\sqrt{4\sqrt{2}-1}} = 2.08... \tag{57}$$

Proof of Lemma 25. We claim that the number $c(\theta)$ of Lemma 24 satisfies

$$\sup_{s\in[0,1]} \frac{\min\left\{1+c(\theta)s^\theta, 2^\theta + c(\theta)(1-s)^\theta\right\}}{(1+s)^\theta} = c(\theta). \tag{58}$$

Indeed, observe that the function $s \mapsto (1+c(\theta)s^\theta)/(1+s)^\theta$ is increasing on $[0,1]$ because one directly computes that its derivative equals $\theta(c(\theta)-s^{1-\theta})/(s^{1-\theta}(1+s)^{1+\theta})$, and by Lemma 24 we have $c(\theta) > 1$ (recall also that $0 < \theta \leqslant 1$). Since the function $s \mapsto (2^\theta + c(\theta)(1-s)^\theta)/(1+s)^\theta$ is decreasing on $[0,1]$, it follows that the supremum that appears in the left hand side of (58) is attained when $1 + c(\theta)s^\theta = 2^\theta + c(\theta)(1-s)^\theta$, or equivalently when $\phi_\theta(s) = (2^\theta-1)/c(\theta)$, where we recall (55). Thus $s = \phi_\theta^{-1}((2^\theta-1)/c(\theta))$ and therefore (58) is equivalent to (56).

Fix $n \in \mathbb{N}$, a stationary reversible Markov chain $\{Z_t\}_{t=0}^\infty$ on $\{1,\ldots,n\}$, and $f:\{1,\ldots,n\} \to X$. For simplicity of notation write $U_t = f(Z_t)$. We shall prove by induction on $t \in \mathbb{N}$ that

$$\mathbb{E}\big[d_X(U_t,U_0)^p\big] \leqslant c(\theta)^p t^{\theta p}\mathbb{E}\big[d_X(U_1,U_0)^p\big]. \tag{59}$$

Lemma 19 shows that (59) holds true if $t = 2^k$ for some $k \in \mathbb{N}\cup\{0\}$ (since $c(\theta) > 1$). So, suppose that $t = (1+s)2^k$ for some $s \in (0,1)$ and $k \in \mathbb{N}\cup\{0\}$. The triangle inequality in L_p, combined with the stationarity of the Markov chain, implies that

$$\begin{aligned}\left(\mathbb{E}\big[d_X(U_t,U_0)^p\big]\right)^{\frac{1}{p}} &\leqslant \left(\mathbb{E}\big[d_X(U_t,U_{2^k})^p\big]\right)^{\frac{1}{p}} + \left(\mathbb{E}\big[d_X(U_{2^k},U_0)^p\big]\right)^{\frac{1}{p}}\\ &= \left(\mathbb{E}\big[d_X(U_{s2^k},U_0)^p\big]\right)^{\frac{1}{p}} + \left(\mathbb{E}\big[d_X(U_{2^k},U_0)^p\big]\right)^{\frac{1}{p}},\end{aligned} \tag{60}$$

and

$$\begin{aligned}\left(\mathbb{E}\big[d_X(U_t,U_0)^p\big]\right)^{\frac{1}{p}} &\leqslant \left(\mathbb{E}\big[d_X(U_t,U_{2^{k+1}})^p\big]\right)^{\frac{1}{p}} + \left(\mathbb{E}\big[d_X(U_{2^{k+1}},U_0)^p\big]\right)^{\frac{1}{p}}\\ &= \left(\mathbb{E}\big[d_X(U_{(1-s)2^k},U_0)^p\big]\right)^{\frac{1}{p}} + \left(\mathbb{E}\big[d_X(U_{2^{k+1}},U_0)^p\big]\right)^{\frac{1}{p}}.\end{aligned} \tag{61}$$

By combining (60) and (61) with Lemma 19 and the inductive hypothesis (59), we see that

$$\frac{\left(\mathbb{E}\big[d_X(U_t,U_0)^p\big]\right)^{\frac{1}{p}}}{\left(\mathbb{E}\big[d_X(U_1,U_0)^p\big]\right)^{\frac{1}{p}}} \leqslant 2^{k\theta}\min\left\{c(\theta)s^\theta+1, c(\theta)(1-s)^\theta+2^\theta\right\} \overset{(58)}{\leqslant} 2^{k\theta}c(\theta)(1+s)^\theta = c(\theta)t^\theta. \qquad \square$$

4. Proof of Proposition 3

Here we justify the validity of Proposition 3 that was stated in the Introduction, thus explaining why we are focusing on quadratic inequalities in the context of the quest for intrinsic characterizations of those metric spaces that admit a bi-Lipschitz embedding into some Alexandrov space that is either nonnegatively or nonpositively curved. The argument below is inspired by the proof of Proposition 15.5.2 in [54].

Proof of Proposition 3. If $c_Y(X) \leqslant D$ for some $(Y,d_Y) \in \mathcal{F}$ then it follows immediately that if $A,B \in M_n(\mathbb{R})$ have nonnegative entries and (Y,d_Y) satisfies the (A,B)-quadratic metric inequality then (X,d_X) satisfies the (A,D^2B)-quadratic metric inequality. The nontrivial direction here is the converse, i.e., suppose that (X,d_X) satisfies the (A,D^2B)-quadratic metric inequality for every two n by n matrices $A,B \in M_n(\mathbb{R})$ with nonnegative entries such that every $(Z,d_Z) \in \mathcal{F}$ satisfies the (A,B)-quadratic metric inequality. The goal is to deduce from this that there exists $(Y,d_Y) \in \mathcal{F}$ for which $c_Y(X) \leqslant D$.

Let $\mathcal{K} \subseteq M_n(\mathbb{R})$ be the set of all n by n matrices $C = (c_{ij})$ for which there exists $(Z,d_Z) \in \mathcal{F}$ and $z_1,\ldots,z_n \in Z$ such that $c_{ij} = d_Z(z_i,z_j)^2$ for every $i,j \in \{1,\ldots,n\}$. Since $\mathcal{F}$ is closed under dilation, we have $[0,\infty)\mathcal{K} \subseteq \mathcal{K}$. Since $\mathcal{F}$ is closed under Pythagorean sums, we have $\mathcal{K}+\mathcal{K} \subseteq \mathcal{K}$. Thus $\mathcal{K}$ is a convex cone.

Write $X = \{x_1,\ldots,x_n\}$. Fix $\varepsilon \in (0,1)$ and suppose for the sake of obtaining a contradiction that there does not exist an embedding of X into any member of $\mathcal{F}$ with distortion less than $D+\varepsilon$. Let $\mathcal{L} \subseteq M_n(\mathbb{R})$ be the set of all n by n symmetric matrices $C = (c_{ij})$ for which there exists $s \in (0,\infty)$ such that $sd_X(i,j)^2 \leqslant c_{ij} \leqslant (D+\varepsilon)^2 sd_X(i,j)^2$ for every $i,j \in \{1,\ldots,n\}$. Our contrapositive assumption means that $\mathcal{K} \cap \mathcal{L} = \emptyset$. Since $\mathcal{K}$ and $\mathcal{L} \cup \{0\}$ are both cones, the separation theorem now implies that there exists a symmetric matrix $H = (h_{ij}) \in M_n(\mathbb{R})$, not all of whose off-diagonal entries vanish, such that

$$\inf_{C \in \mathcal{L}} \sum_{i=1}^n \sum_{j=1}^n h_{ij} c_{ij} \geqslant 0 \geqslant \sup_{C \in \mathcal{K}} \sum_{i=1}^n \sum_{j=1}^n h_{ij} c_{ij}. \qquad (62)$$

Define $A,B \in M_n(\mathbb{R})$ by setting for every $i,j \in \{1,\ldots,n\}$,

$$a_{ij} \stackrel{\mathrm{def}}{=} \begin{cases} h_{ij} & \text{if } h_{ij} \geqslant 0, \\ 0 & \text{if } h_{ij} < 0, \end{cases} \qquad \text{and} \qquad b_{ij} \stackrel{\mathrm{def}}{=} \begin{cases} |h_{ij}| & \text{if } h_{ij} < 0, \\ 0 & \text{if } h_{ij} \geqslant 0, \end{cases}$$

The right hand inequality in (62), combined with the definition of $\mathcal{K}$, implies that every $(Y,d_Y) \in \mathcal{F}$ satisfies the (A,B)-quadratic metric inequality. By our assumption on X, this implies that

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij} d_X(x_i,x_j)^2 \leqslant D^2 \sum_{i=1}^n \sum_{j=1}^n b_{ij} d_X(x_i,x_j)^2 < (D+\varepsilon)^2 \sum_{i=1}^n \sum_{j=1}^n b_{ij} d_X(x_i,x_j)^2, \qquad (63)$$

where we used the fact that not all the off-diagonal entries of H vanish, so all the sums appearing (63) are positive. Consequently, if we set

$$\forall\, i,j \in \{1,\ldots,n\}, \qquad c_{ij} \stackrel{\mathrm{def}}{=} \begin{cases} (D+\varepsilon)^2 d_X(x_i,x_j)^2 & \text{if } h_{ij} < 0 \\ d_X(x_i,x_j)^2 & \text{if } h_{ij} \geqslant 0, \end{cases}$$

then $C = (c_{ij}) \in \mathcal{L}$ and by (63) we have $\sum_{i=1}^n \sum_{j=1}^n h_{ij} c_{ij} < 0$. This contradicts the left hand inequality in (62). $\square$

5. Subsets of Hadamard spaces

As we discussed in the introduction, it is a major open problem to characterize those finite metric spaces that admit a bi-Lipschitz (or even isometric) embedding into some Hadamard space. By Proposition 3, this amounts to understanding those quadratic metric inequalities that hold true in any Hadamard space. In this section we shall derive potential families of such inequalities.

An equivalent characterization of when a metric space (X, d_X) is a Hadamard space is the requirement that there exists a mapping $\mathfrak{B}$ that assigns a point $\mathfrak{B}(\mu) \in X$ to every finitely supported probability measure μ on X with the property that $\mathfrak{B}(\delta_x) = x$ for every $x \in X$ (i.e., $\mathfrak{B}$ is a barycenter map) and every finitely supported probability measure μ on X satisfies the following inequality for every $x \in X$.

$$d_X(x, \mathfrak{B}(\mu))^2 + \int_X d_X(\mathfrak{B}(\mu), y)^2 \mathrm{d}\mu(y) \leqslant \int_X d_X(x, y)^2 \mathrm{d}\mu(y). \tag{64}$$

For the proof that (X, d_X) is a Hadamard space if and only if it satisfies (64), see e.g. Lemma 4.4. and Theorem 4.9 in [88]. One could extend the validity of (64) to probability measures that are not necessarily finitely supported, but this will be irrelevant for our purposes.

Lemma 27 below yields a general recipe for producing quadratic metric inequalities that hold true in any Hadamard space.

Lemma 27. *Fix $n \in \mathbb{N}$ and $p_1, \ldots, p_n, q_1, \ldots, q_n \in (0, 1)$ such that $\sum_{i=1}^n p_i = \sum_{j=1}^n q_j = 1$. Suppose that $A = (a_{ij}), B = (b_{ij}) \in M_n(\mathbb{R})$ are n by n matrices with nonnegative entries that satisfy*

$$\forall\, i, j \in \{1, \ldots, n\}, \qquad \sum_{k=1}^n a_{ik} + \sum_{k=1}^n b_{kj} = p_i + q_j. \tag{65}$$

If (X, d_X) is a Hadamard space then for every $x_1, \ldots, x_n \in X$ we have

$$\sum_{i=1}^n \sum_{j=1}^n \frac{a_{ij} b_{ij}}{a_{ij} + b_{ij}} d_X(x_i, x_j)^2 \leqslant \sum_{i=1}^n \sum_{j=1}^n p_i q_j d_X(x_i, x_j)^2. \tag{66}$$

Proof. Writing $z = \mathfrak{B}\Big(\sum_{i=1}^n p_i \delta_{x_i} \Big)$, by (64) for every $j \in \{1, \ldots, n\}$ we have

$$d_X(x_j, z)^2 + \sum_{i=1}^n p_i d_X(x_i, z)^2 \leqslant \sum_{i=1}^n p_i d_X(x_i, x_j)^2. \tag{67}$$

By multiplying (67) by q_j and summing over $j \in \{1, \ldots, n\}$ we get

$$\sum_{j=1}^n q_j d_X(x_j, z)^2 + \sum_{i=1}^n p_i d_X(x_i, z)^2 \leqslant \sum_{i=1}^n \sum_{j=1}^n p_i q_j d_X(x_i, x_j)^2. \tag{68}$$

Hence,

$$\sum_{j=1}^n q_j d_X(x_j, z)^2 + \sum_{i=1}^n p_i d_X(x_i, z)^2 = \sum_{i=1}^n \sum_{j=1}^n \big(a_{ij} d_X(x_i, z)^2 + b_{ij} d_X(x_j, z)^2 \big) \tag{69}$$

$$\geqslant \sum_{i=1}^n \sum_{j=1}^n \frac{a_{ij} b_{ij}}{a_{ij} + b_{ij}} d_X(x_i, x_j)^2, \tag{70}$$

where in (69) we used (65), and (70) holds true because $d_X(x_i,z)+d_X(x_j,z) \geqslant d_X(x_i,x_j)$ for every $i,j \in \{1,\ldots,n\}$, and for every $s,t,\gamma \in [0,\infty)$ we have (by e.g. Cauchy-Schwarz),

$$\min_{\substack{\alpha,\beta\in[0,\infty)\\ \alpha+\beta\geqslant\gamma}} \left(s\alpha^2+t\beta^2\right) = \frac{st\gamma^2}{s+t}. \tag{71}$$

The desired estimate (66) is a combination of (68) and (70). □

The proof of Lemma 27 is a systematic way to exploit the existence of barycenters in order to deduce quadratic metric inequalities, under the crucial constraint that the final inequality is allowed to involve only distances within the subset $\{x_1,\ldots,x_n\} \subseteq X$. The barycentric inequality (64) is used in (67), but one must then remove all reference to the auxiliary point z since it need not be part of the given subset $\{x_1,\ldots,x_n\}$. It is natural to do so by incorporating the triangle inequality $d_X(x_i,z)+d_X(x_j,z) \geqslant d_X(x_i,x_j)$ for some $i,j \in \{1,\ldots,n\}$. This inequality is distributed among the possible pairs $i,j \in \{1,\ldots,n\}$ through a general choice of re-weighting matrices A,B, with the final step in (70) being sharp due to (71). A more general scheme along these lines will be described in Section 5.2 below, but (an iterative applications of) the above simple scheme is already powerful, and in fact we do not know whether or not it yields a characterization of subsets of Hadamard spaces; see Question 31 below.

A notable special case of Lemma 27 is when $(p_1,\ldots,p_n)=(q_1,\ldots,q_n)$ and there exists a permutation $\pi \in S_n$ such that $a_{i\pi(i)} = b_{\pi^{-1}(i)i} = p_i$ for every $i \in \{1,\ldots,n\}$, while all the other entries of the matrices A and B vanish. In this case one arrives at the following useful inequality.

Corollary 28. *Suppose that (X,d_X) is a Hadamard space. Then for every $n\in\mathbb{N}$, every $x_1,\ldots,x_n$, every $p_1,\ldots,p_n \in [0,1]$ with $\sum_{j=1}^n p_j = 1$ and every permutation $\pi \in S_n$ we have*

$$\sum_{i=1}^n \frac{p_i p_{\pi(i)}}{p_i+p_{\pi(i)}} d_X(x_i,x_{\pi(i)})^2 \leqslant \sum_{i=1}^n\sum_{j=1}^n p_ip_j d_X(x_i,x_j)^2. \tag{72}$$

When $n=4$ and $\pi=(1,3)(2,4)$, Corollary 28 becomes

Corollary 29. *Suppose that (X,d_X) be a Hadamard space and fix $x_1,x_2,x_3,x_4 \in X$. Then for every $p_1,p_2,p_3,p_4 \in [0,\infty)$ we have*

$$\begin{aligned} p_1p_2d_X(x_1,x_2)^2 + p_2p_3d_X(x_2,x_3)^2 + p_3p_4d_X(x_3,x_4)^2 + p_4p_1d_X(x_4,x_1)^2 \\ \geqslant \frac{p_1p_3(p_2+p_4)}{p_1+p_3} d_X(x_1,x_3)^2 + \frac{p_2p_4(p_1+p_3)}{p_2+p_4} d_X(x_2,x_4)^2. \end{aligned} \tag{73}$$

To pass from (72) to (73) note that (73) is homogeneous of order 2 in (p_1,p_2,p_3,p_4), so we may assume that $p_1+p_2+p_3+p_4=1$. Now (73) is a direct application of (72) with the above specific choice of permutation π, while subtracting from both sides of (72) those multiples of $d_X(x_1,x_3)^2$ and $d_X(x_2,x_4)^2$ that appear in the right hand side of (72).

When $p_1+p_3=p_2+p_4=1$, Corollary 29 becomes Sturm's weighted quadruple inequality [88], which asserts that for every Hadamard space (X,d_X), every $x_1,x_2,x_3,x_4 \in X$, and every $s,t \in [0,1]$,

$$\begin{aligned} s(1-s)d_X(x_1,x_3)^2 + t(1-t)d_X(x_2,x_4)^2 \\ \leqslant std_X(x_1,x_2)^2 + (1-s)td_X(x_2,x_3)^2 + (1-s)(1-t)d_X(x_3,x_4)^2 + s(1-t)d_X(x_4,x_1)^2. \end{aligned} \tag{74}$$

As explained in [88, Proposition 2.4], by choosing the parameters s,t appropriately in (74) one obtains an important quadruple comparison inequality of Reshetnyak [84] (see also [36] or [44, Lemma 2.1]), asserting that for every Hadamard space (X,d_X) and every $x_1,x_2,x_3,x_4 \in X$,

$$d_X(x_1,x_3)^2 + d_X(x_2,x_4)^2 \leqslant d_X(x_1,x_2)^2 + d_X(x_2,x_3)^2 + 2d_X(x_3,x_4)d_X(x_4,x_1). \tag{75}$$

The coefficients in (73) have 3 degrees of freedom while in (74) they have 2 degrees of freedom. This additional flexibility yields a proof of the validity of the Ptolemy inequality (4) in Hadamard spaces. The fact that the Ptolemy inequality holds true in Hadamard spaces was proved in [37, 25], and an alternative proof was given in [17]. Both of these proofs rely on comparisons with ideal configurations in the Euclidean plane (see [15, §II.1]), combined with the classical Ptolemy theorem in Euclidean geometry. Corollary 30 below shows how the Ptolemy inequality is a direct consequence of (73), thus yielding an intrinsic proof that does not proceed through an embedding argument.

Corollary 30. *Let (X,d_X) be a Hadamard space and $x_1,x_2,x_3,x_4\in X$. Write $d_{ij}=d_X(x_i,x_j)$ for every $i,j\in\{1,\ldots,n\}$. Then*

$$d_{12}d_{34}+d_{23}d_{41}-d_{13}d_{24}\geqslant \frac{\big((d_{12}d_{23}+d_{34}d_{41})\,d_{13}-(d_{12}d_{41}+d_{23}d_{34})\,d_{24}\big)^2}{2\,(d_{12}d_{41}+d_{23}d_{34})\,(d_{12}d_{23}+d_{34}d_{41})}\geqslant 0. \tag{76}$$

Proof. The proof of (76) is nothing more than an application of Corollary 29 with the following specific choices of $p_1,p_2,p_3,p_4\in[0,\infty)$.

$$p_1\stackrel{\mathrm{def}}{=}\frac{d_{34}}{d_{41}}\cdot\frac{d_{23}+d_{41}}{d_{12}+d_{34}},\quad p_2\stackrel{\mathrm{def}}{=}\frac{d_{41}}{d_{12}}\cdot\frac{d_{12}+d_{34}}{d_{23}+d_{41}},\quad p_3\stackrel{\mathrm{def}}{=}\frac{d_{12}}{d_{23}}\cdot\frac{d_{23}+d_{41}}{d_{12}+d_{34}},\quad p_4\stackrel{\mathrm{def}}{=}\frac{d_{23}}{d_{34}}\cdot\frac{d_{12}+d_{34}}{d_{23}+d_{41}}.$$

A substitution of these values into (73) yields

$$\begin{aligned}2d_{12}d_{34}+2d_{23}d_{41}&\geqslant\frac{d_{12}d_{23}+d_{34}d_{41}}{d_{12}d_{41}+d_{23}d_{34}}d_{13}^2+\frac{d_{12}d_{41}+d_{23}d_{34}}{d_{12}d_{23}+d_{34}d_{41}}d_{24}^2\\&=2d_{13}d_{24}+\frac{((d_{12}d_{23}+d_{34}d_{41})\,d_{13}-(d_{12}d_{41}+d_{23}d_{34})\,d_{24})^2}{(d_{12}d_{41}+d_{23}d_{34})\,(d_{12}d_{23}+d_{34}d_{41})}.\end{aligned}$$

□

5.1. **Iterative applications of Lemma 27.** The case $s=t=1/2$ of (73) becomes the roundness 2 inequality (3), i.e., for every Hadamard space (X,d_X) and every $x_1,x_2,x_3,x_4\in X$ we have

$$d_X(x_1,x_3)^2+d_X(x_2,x_4)^2\leqslant d_X(x_1,x_2)^2+d_X(x_2,x_3)^2+d_X(x_3,x_4)^2+d_X(x_4,x_1)^2. \tag{77}$$

In [24], Enflo iterated (77) (while exploiting cancellations) so as to yield the following inequality, which holds for every Hadamard space (X,d_X), every $n\in\mathbb{N}$ and every $f:\{-1,1\}^n\to X$.

$$\sum_{x\in\{-1,1\}^n} d_X\left(f(x),f(-x)\right)^2\leqslant\sum_{i=1}^n\sum_{x\in\{-1,1\}^n} d_X\left(f(x),f(x_1,\ldots,x_{i-1},-x_i,x_{i+1},\ldots,x_n)\right)^2. \tag{78}$$

In today's terminology (78) says that every Hadamard space has Enflo type 2 with constant 1 (see also [77]). The argument in [49] yields a different iterative application of (77) (again, exploiting cancellations via a telescoping argument), showing that mappings from the iterated diamond graph (see [74]) into any Hadamard space satisfy a certain quadratic metric inequality. Similar reasoning (as in [48]) yields a quadratic metric inequality for Hadamard space-valued mappings on the Laakso graphs (see [41, 45]). The value of the above iterative applications of (77) is that they yield inequalities on metric spaces of unbounded cardinality (hypercubes, diamond graphs, Laakso graphs) that serve as obstructions to bi-Lipschitz embeddings of these spaces into any Hadamard space: these inequalities imply that any such embedding must incur distortion that tends to ∞ as the size of the underlying space tends to ∞ (in fact, these inequalities yield sharp bounds).

We therefore see that by applying Lemma 27 multiple times one could obtain quadratic metric inequalities that yield severe restrictions on those metric spaces that admit a bi-Lipschits embedding into some Hadamard space. Specifically, one could apply Lemma 27 to several configurations of points and several choices of weights, and consider a weighted average of the resulting inequalities. This yields the estimate

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{m}\frac{c_k a_{ij}^k b_{ij}^k}{a_{ij}^k+b_{ij}^k}d_X(x_i,x_j)^2\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{m}c_k p_i^k q_j^k d_X(x_i,x_j)^2, \tag{79}$$

which is valid for every Hadamard space (X,d_X), every $m,n\in\mathbb{N}$, every $x_1,\ldots,x_n\in X$, every $\{c_k\}_{k=1}^m\subseteq(0,\infty)$, every $\left\{p_i^k,q_i^k:\ i\in\{1,\ldots,n\},\ k\in\{1,\ldots,m\}\right\}\subseteq(0,\infty)$ with

$$\sum_{i=1}^{n}p_i^k=\sum_{j=1}^{n}q_j^k=1, \tag{80}$$

and every choice of n by n matrices $\{A_k=(a_{ij}^k)\}_{k=1}^m,\{B_k=(b_{ij}^k)\}_{k=1}^m\subseteq M_n(\mathbb{R})$ with nonnegative entries, such that for every $i,j\in\{1,\ldots,n\}$ and $k\in\{1,\ldots,m\}$,

$$\sum_{s=1}^{n}a_{is}^k+\sum_{s=1}^{n}b_{sj}^k=p_i^k+q_j^k. \tag{81}$$

By collecting terms in (79) so that for every $i,j\in\{1,\ldots,n\}$ no multiple of $d_X(x_i,x_j)^2$ appears in both sides of the inequality, as was done in e.g. (78), one arrives at the following estimate.

$$\begin{aligned}\sum_{\substack{i,j\in\{1,\ldots,n\}\\ \sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)>0}}\sum_{k=1}^{m}c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)d_X(x_i,x_j)^2\\ \leqslant \sum_{\substack{i,j\in\{1,\ldots,n\}\\ \sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)<0}}\sum_{k=1}^{m}c_k\left(p_i^kq_j^k-\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}\right)d_X(x_i,x_j)^2.\end{aligned} \tag{82}$$

To the best of our knowledge, all of the previously used quadratic metric inequalities on general Hadamard spaces are of the form (82). We therefore ask whether the inequalities of the form (82) capture the totality of those quadratic metric inequalities that are valid in Hadamard spaces.

Question 31. Is it true that for every $D\in[1,\infty)$ there exists some $c(D)\in[1,\infty)$ such that a metric space (X,d_X) embeds with distortion at most $c(D)$ into some Hadamard space provided

$$\begin{aligned}\sum_{\substack{i,j\in\{1,\ldots,n\}\\ \sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)>0}}\sum_{k=1}^{m}c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)d_X(x_i,x_j)^2\\ \leqslant D^2\cdot \sum_{\substack{i,j\in\{1,\ldots,n\}\\ \sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)<0}}\sum_{k=1}^{m}c_k\left(p_i^kq_j^k-\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}\right)d_X(x_i,x_j)^2,\end{aligned}$$

for all $m,n\in\mathbb{N}$, all $c_k,p_i^k,q_i^k,a_{ij}^k,b_{ij}^k\in[0,\infty)$ satisfying (80) and (81), and all $x_1,\ldots,x_n\in X$?

Recall that there are useful metric inequalities, which are not quadratic metric inequalities, that hold true in any Hadamard space, such as Reshetnyak's inequality (75) or the Ptolemy inequality (4). However, we already know through Proposition 3 that quadratic metric inequalities fully characterize subsets of Hadamard spaces. And, in the case of Reshetnyak's inequality or the Ptolemy inequality, we have seen above how to deduce them explicitly from a quadratic metric inequality (the key point to note here is that the various coefficients that appear in (82) can be optimized so as to depend on the distances $\{d_X(x_i,x_j)\}_{i,j\in\{1,\ldots,n\}}$).

A negative answer to Question 31 would be very interesting, as it would yield a new family of metric spaces that fail to admit a bi-Lipschitz embedding into any Hadamard space, and correspondingly a new family of quadratic metric inequalities which hold true in any Hadamard space yet do not follow from the above procedure for obtaining such inequalities.

As discussed in the Introduction, it is not known whether or not for every metric space (X,d_X) there exists a Hadamard space (Y,d_Y) with $c_Y(X,\sqrt{d_X})<\infty$. If this were true then Question 32 below would have a positive answer. Conversely, a positive answer to both Question 31 and Question 32 would imply that the 1/2-snowflake of any metric space admits a bi-Lipschitz embedding into some Hadamard space.

Question 32. Is it true that every metric space (X,d_X) satisfies

$$\sum_{\substack{i,j\in\{1,\ldots,n\}\\ \sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)>0}}\sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)d_X(x_i,x_j)$$
$$\lesssim \sum_{\substack{i,j\in\{1,\ldots,n\}\\ \sum_{k=1}^m c_k\left(\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}-p_i^kq_j^k\right)<0}}\sum_{k=1}^m c_k\left(p_i^kq_j^k-\frac{a_{ij}^kb_{ij}^k}{a_{ij}^k+b_{ij}^k}\right)d_X(x_i,x_j),$$

for all $m,n\in\mathbb{N}$, all $c_k,p_i^k,q_i^k,a_{ij}^k,b_{ij}^k\in[0,\infty)$ satisfying (80) and (81), and all $x_1,\ldots,x_n\in X$?

Question 32 seems tractable, but at present we do not know whether or not its answer is positive. A negative answer to Question 32 would yield for the first time a metric space (X,d_X) such that $(X,\sqrt{d_X})$ fails to admit a bi-Lipschitz embedding into any Hadamard space, in sharp contrast to the case of embeddings into Alexandrov spaces of nonnegative curvature. In the same vein, a proof that every Hadamard space admits a sequence of bounded degree expanders would resolve Question 32 negatively. It is true that inequality (73) is not an obstruction to the validity of Question 32, i.e., for every metric space (X,d_X), every $p_1,p_2,p_3,p_4\in[0,\infty)$ and every $x_1,x_2,x_3,x_4\in X$ we have

$$p_1p_2d_X(x_1,x_2)+p_2p_3d_X(x_2,x_3)+p_3p_4d_X(x_3,x_4)+p_4p_1d_X(x_4,x_1)$$
$$\geqslant \frac{p_1p_3(p_2+p_4)}{p_1+p_3}d_X(x_1,x_3)+\frac{p_2p_4(p_1+p_3)}{p_2+p_4}d_X(x_2,x_4). \quad (83)$$

Also (actually, as a consequence of (83)), Reshetnyak's inequality and the Ptolemy inequality hold true in any square root of a metric space, i.e., for every metric space (X,d_X) and $x_1,x_2,x_3,x_4\in X$,

$$d_X(x_1,x_3)+d_X(x_2,x_4)\leqslant d_X(x_1,x_2)+d_X(x_2,x_3)+2\sqrt{d_X(x_3,x_4)d_X(x_4,x_1)},$$

and

$$\sqrt{d_X(x_1,x_3)d_X(x_2,x_4)}\leqslant\sqrt{d_X(x_1,x_2)d_X(x_3,x_4)}+\sqrt{d_X(x_2,x_3)d_X(x_4,x_1)}.$$

It is possible (and instructive) to prove these inequalities while using only the triangle inequality, but this seems to require a somewhat tedious case analysis. Alternatively, one could verify (83) by using the fact that the square root of any four-point metric space admits an isometric embedding into a Hilbert space; see e.g. [20, Proposition 2.6.2].

Lemma 33 below asserts that the conclusion of Lemma 27 holds true in any square-root of a metric space, with a loss of a constant factor. This is a special case of Question 32 that falls sort of a positive answer in general due to the fact that we want to iterate the resulting inequality, in which case the constant factor loss could accumulate.

Lemma 33. *Fix $n \in \mathbb{N}$ and $p_1, \ldots, p_n, q_1, \ldots, q_n \in (0,1)$ such that $\sum_{i=1}^n p_i = \sum_{j=1}^n q_j = 1$. Suppose that $A = (a_{ij}), B = (b_{ij}) \in M_n(\mathbb{R})$ are n by n matrices with nonnegative entries that satisfy (65). Then for every metric space (X, d_X) and every $x_1, \ldots, x_n \in X$ we have*

$$\sum_{i=1}^n \sum_{j=1}^n \frac{a_{ij} b_{ij}}{a_{ij} + b_{ij}} d_X(x_i, x_j) \leqslant 3 \sum_{i=1}^n \sum_{j=1}^n p_i q_j d_X(x_i, x_j).$$

Proof. Let $F : \{x_1, \ldots, x_n\} \to \ell_\infty$ be any isometric embedding of the metric space $(\{x_1, \ldots, x_n\}, d_X)$ into ℓ_∞. By convexity we have

$$\begin{aligned}\sum_{i=1}^n p_i \Big\| F(x_i) - \sum_{j=1}^n q_j F(x_j) \Big\|_\infty &= \sum_{i=1}^n p_i \Big\| \sum_{j=1}^n q_j (F(x_i) - F(x_j)) \Big\|_\infty \\ &\leqslant \sum_{i=1}^n \sum_{j=1}^n p_i q_j \|F(x_i) - F(x_j)\|_\infty = \sum_{i=1}^n \sum_{j=1}^n p_i q_j d_X(x_i, x_j),\end{aligned}$$

and similarly,

$$\begin{aligned}\sum_{i=1}^n q_i \Big\| F(x_i) - \sum_{j=1}^n q_j F(x_j) \Big\|_\infty &\leqslant \sum_{i=1}^n q_i \Big\| F(x_i) - \sum_{j=1}^n p_j F(x_j) \Big\|_\infty + \Big\| \sum_{i=1}^n \sum_{j=1}^n p_i q_j (F(x_i) - F(x_j)) \Big\|_\infty \\ &\leqslant 2 \sum_{i=1}^n \sum_{j=1}^n p_i q_j \|F(x_i) - F(x_j)\|_\infty = 2 \sum_{i=1}^n \sum_{j=1}^n p_i q_j d_X(x_i, x_j).\end{aligned}$$

So, if we denote $z \stackrel{\mathrm{def}}{=} \sum_{k=1}^n q_k F(x_k)$ then

$$\begin{aligned}3 \sum_{i=1}^n \sum_{j=1}^n p_i q_j d_X(x_i, x_j) &\geqslant \sum_{i=1}^n p_i \|F(x_i) - z\|_\infty + \sum_{j=1}^n q_j \|F(x_j) - z\|_\infty \\ &\stackrel{(65)}{=} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ij} \|F(x_i) - z\|_\infty + b_{ij} \|F(x_j) - z\|_\infty \right) \\ &\geqslant \sum_{i=1}^n \sum_{j=1}^n \min\{a_{ij}, b_{ij}\} \left(\|F(x_i) - z\|_\infty + \|F(x_j) - z\|_\infty \right) \\ &\geqslant \sum_{i=1}^n \sum_{j=1}^n \frac{a_{ij} b_{ij}}{a_{ij} + b_{ij}} \|F(x_i) - F(x_j)\|_\infty \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{a_{ij} b_{ij}}{a_{ij} + b_{ij}} d_X(x_i, x_j).\end{aligned}$$

□

5.2. **A hierarchy of quadratic metric inequalities.** The quadratic metric inequalities of Section 5.1 are part of a first level of a hierarchy of quadratic metric inequalities that hold true in any Hadamard space. We shall now describe these inequalities, which quickly become quite complicated and unwieldy. We conjecture that the entire hierarchy of inequalities thus obtained characterizes subsets of Hadamard spaces; see Question 34 below. Due to the generality of these inequalities, this conjecture could be quite tractable. But, even if it has a positive answer then it would yield a complicated, and therefore perhaps less useful, characterization of subsets of Hadamard spaces, and it would still be very interesting to find a smaller family of inequalities that characterizes subsets of Hadamard spaces, in the spirit of Question 31.

Let (X,d_X) be a Hadamard space. The barycentric inequality (64) has the following counterpart as a formal consequence, which is an inequality that allows one to control the distance between barycenters of two probability measures. Let μ,ν be finitely supported probability measures on X. By applying (64) twice we see that

$$d_X(\mathfrak{B}(\nu),\mathfrak{B}(\mu))^2+\int_X d_X(\mathfrak{B}(\mu),x)^2\mathrm{d}\mu(x)\leqslant \int_X d_X(\mathfrak{B}(\nu),x)^2\mathrm{d}\mu(x)$$
$$\leqslant \int_X\left(\int_X d_X(x,y)^2\mathrm{d}\nu(y)-\int_X d_X(\mathfrak{B}(\nu),y)^2\mathrm{d}\nu(y)\right)\mathrm{d}\mu(x).$$

Thus

$$d_X(\mathfrak{B}(\nu),\mathfrak{B}(\mu))^2+\int_X d_X(\mathfrak{B}(\mu),x)^2\mathrm{d}\mu(x)+\int_X d_X(\mathfrak{B}(\nu),y)^2\mathrm{d}\nu(y)$$
$$\leqslant \iint_{X\times X} d_X(x,y)^2\mathrm{d}\mu(x)\mathrm{d}\nu(y). \quad (84)$$

Both (64) and (84) will be used repeatedly in what follows.

5.2.1. *An inductive construction.* Fix $n\in\mathbb{N}$ and $x_1,\ldots,x_n\in X$. Fix also a sequence of integers $\{m_s\}_{s=0}^\infty\subseteq\mathbb{N}$ with $m_0=n$. Suppose that we are given $\mu_{s,a}^{k+1,b}\in[0,\infty)$ for every $k\in\mathbb{N}\cup\{0\}$, $s\in\{0,\ldots,k\}$, $a\in\{1,\ldots,m_s\}$ and $b\in\{1,\ldots,m_{k+1}\}$, such that

$$\forall\, b\in\{1,\ldots,m_{k+1}\},\qquad \sum_{s=0}^{k}\sum_{a=1}^{m_s}\mu_{s,a}^{k+1,b}=1.$$

We shall now proceed to define by induction on $k\in\mathbb{N}\cup\{0\}$ auxiliary points $x_a^s\in X$ for every $s\in\{0,\ldots,k\}$ and $a\in\{1,\ldots,m_s\}$. Our construction will also yield for every $i,j\in\{1,\ldots,n\}$, $s,t,\sigma,\tau\in\{0,\ldots,k\}$, $a\in\{1,\ldots,m_s\}$, $\alpha\in\{1,\ldots,m_\sigma\}$, $b\in\{1,\ldots,m_t\}$ and $\beta\in\{1,\ldots,m_\tau\}$ nonnegative weights $U_{\sigma,\tau,\alpha,\beta}^{s,t,a,b},V_{i,j}^{s,t,a,b}\in[0,\infty)$ that satisfy the inequality

$$d_X(x_a^s,x_b^t)^2+\sum_{\sigma=0}^{k}\sum_{\tau=0}^{k}\sum_{\alpha=1}^{m_\sigma}\sum_{\beta=1}^{m_\tau}U_{\sigma,\tau,\alpha,\beta}^{s,t,a,b}d_X(x_\alpha^\sigma,x_\beta^\tau)^2\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}V_{i,j}^{s,t,a,b}d_X(x_i,x_j)^2. \quad (85)$$

The induction starts by setting $x_a^0=x_a$ for $a\in\{1,\ldots,n\}$. Also, for every $a,b,\alpha,\beta\in\{1,\ldots,n\}$ set $U_{0,0,\alpha,\beta}^{0,0,a,b}=0$ and $V_{\alpha,\beta}^{0,0,a,b}=\mathbf{1}_{\{(\alpha,\beta)=(a,b)\}}$, thus satisfying (85) vacuously.

Suppose now that we have defined $x_a^s\in X$ for every $s\in\{0,\ldots,k\}$ and $a\in\{1,\ldots,m_s\}$. Consider the probability measures

$$\forall\, b\in\{1,\ldots,m_{k+1}\},\qquad \mu^{k+1,b}\stackrel{\mathrm{def}}{=}\sum_{s=0}^{k}\sum_{a=1}^{m_s}\mu_{s,a}^{k+1,b}\delta_{x_a^s},$$

and define

$$\forall\, b\in\{1,\ldots,m_{k+1}\},\qquad x_b^{k+1}\stackrel{\mathrm{def}}{=}\mathfrak{B}\left(\mu^{k+1,b}\right).$$

Suppose that $s \in \{0,\ldots,k\}$, $a \in \{1,\ldots,m_s\}$ and $b \in \{1,\ldots,m_{k+1}\}$. Then by (64) we have

$$d_X(x_a^s, x_b^{k+1})^2 + \sum_{\tau=0}^{k}\sum_{\beta=1}^{m_\tau} \mu_{\tau,\beta}^{k+1,b} d_X(x_b^{k+1}, x_\beta^\tau)^2 \leqslant \sum_{t=0}^{k}\sum_{c=1}^{m_t} \mu_{t,c}^{k+1,b} d_X(x_a^s, x_c^t)^2.$$

In combination with the inductive hypothesis (85), this implies that the desired estimate (85) would also hold true when $|\{s,t\} \cap \{k+1\}| = 1$ once we introduce the following inductive definitions.

$$U_{\sigma,\tau,\beta,\alpha}^{k+1,s,b,a} = U_{\sigma,\tau,\alpha,\beta}^{s,k+1,a,b} \stackrel{\mathrm{def}}{=} \mathbf{1}_{\{(\sigma,\alpha,\tau)\in\{k+1\}\times\{b\}\times\{0,\ldots,k\}\}} \mu_{\tau,\beta}^{k+1,b} + \mathbf{1}_{\{\{\sigma,\tau\}\subseteq\{0,\ldots,k\}\}} \sum_{t=0}^{k}\sum_{c=1}^{m_t} \mu_{t,c}^{k+1,b} U_{\sigma,\tau,\alpha,\beta}^{s,t,a,c},$$

and

$$V_{i,j}^{s,k+1,a,b} \stackrel{\mathrm{def}}{=} \sum_{t=0}^{k}\sum_{c=1}^{m_t} \mu_{t,c}^{k+1,b} V_{i,j}^{s,t,a,c}.$$

It remains to ensure the validity of (85) when $s = t = k+1$. So, fix $a, b \in \{1,\ldots,k+1\}$ and apply (84) so as to obtain the estimate

$$d_X(x_a^{k+1}, x_b^{k+1})^2 + \sum_{\tau=0}^{k}\sum_{\beta=1}^{m_\tau} \mu_{\tau,\beta}^{k+1,a} d_X(x_a^{k+1}, x_\beta^\tau)^2 + \sum_{\tau=0}^{k}\sum_{\beta=1}^{m_\tau} \mu_{\tau,\beta}^{k+1,b} d_X(x_b^{k+1}, x_\beta^\tau)^2 \leqslant \sum_{t=0}^{k}\sum_{\theta=0}^{k}\sum_{p=1}^{m_t}\sum_{q=1}^{m_\theta} \mu_{t,p}^{k+1,a}\mu_{\theta,q}^{k+1,b} d_X(x_p^t, x_q^\theta)^2.$$

In combination with the inductive hypothesis (85), this implies that the desired estimate (85) would also hold true when $s = t = k+1$ once we introduce the following inductive definitions.

$$U_{\sigma,\tau,\alpha,\beta}^{k+1,k+1,a,b} \stackrel{\mathrm{def}}{=} \mathbf{1}_{\{(\sigma,\alpha,\tau)\in\{k+1\}\times\{a,b\}\times\{0,\ldots,k\}\}} \mu_{\tau,\beta}^{k+1,\alpha} + \mathbf{1}_{\{\{\sigma,\tau\}\subseteq\{0,\ldots,k\}\}} \sum_{t=0}^{k}\sum_{\theta=0}^{k}\sum_{p=1}^{m_t}\sum_{q=1}^{m_\theta} \mu_{t,p}^{k+1,a}\mu_{\theta,q}^{k+1,b} U_{\sigma,\tau,\alpha,\beta}^{t,\theta,p,q},$$

and

$$V_{i,j}^{k+1,k+1,a,b} = \sum_{t=0}^{k}\sum_{\theta=0}^{k}\sum_{p=1}^{m_t}\sum_{q=1}^{m_\theta} \mu_{t,p}^{k+1,a}\mu_{\theta,q}^{k+1,b} V_{i,j}^{t,\theta,p,q}.$$

This concludes our inductive construction of auxiliary points, which satisfy the inequality (85). We shall now show how to remove the auxiliary points so as to obtain bona fide quadratic metric inequalities that involve only points from the subset $\{x_1,\ldots,x_n\} \subseteq X$.

5.2.2. *Deriving quadratic metric inequalities.* Suppose that for every $s,t \in \{0,\ldots,k\}$, $a \in \{0,\ldots,m_s\}$ and $b \in \{0,\ldots,m_t\}$ we are given a nonnegative weight $\Gamma_{a,b}^{s,t} \in [0,\infty)$. By multiplying (85) by $\Gamma_{a,b}^{s,t}$ and summing the resulting inequalities, we obtain the estimate

$$\sum_{s=0}^{k}\sum_{t=0}^{k}\sum_{a=1}^{m_s}\sum_{b=1}^{m_t} E_{a,b}^{s,t} d_X(x_a^s, x_b^t)^2 \leqslant \sum_{i=1}^{n}\sum_{j=1}^{n} F_{i,j} d_X(x_i, x_j)^2, \tag{86}$$

where

$$E_{a,b}^{s,t} \stackrel{\mathrm{def}}{=} \Gamma_{a,b}^{s,t} + \sum_{\sigma=0}^{k}\sum_{\tau=0}^{k}\sum_{\alpha=1}^{m_\sigma}\sum_{\beta=1}^{m_\tau} \Gamma_{\alpha,\beta}^{\sigma,\tau} U_{s,t,a,b}^{\sigma,\tau,\alpha,\beta},$$

and

$$F_{i,j} \stackrel{\mathrm{def}}{=} \sum_{s=0}^{k}\sum_{t=0}^{k}\sum_{a=1}^{m_s}\sum_{b=1}^{m_t}\Gamma_{a,b}^{s,t}V_{ij}^{s,t,a,b}.$$

Denote

$$\mathsf{S}_k \stackrel{\mathrm{def}}{=} \left\{x_a^s:\ s\in\{0,\ldots,k\} \text{ and } a\in\{1,\ldots,m_s\}\right\}\subseteq X.$$

Any $\zeta\in\bigcup_{\ell=1}^{\infty}\mathsf{S}_k^{\ell}$ will be called below a path in S_k. If $\zeta=(\zeta_0,\ldots,\zeta_\ell)$ for some $\ell\in\mathbb{N}$ then we write $\ell(\zeta)=\ell$. The points $\zeta_0,\zeta_{\ell(\zeta)}$ are called the endpoints of the path ζ. The path ζ is called non-repetitive if the points $\zeta_0,\ldots,\zeta_{\ell(\zeta)}$ are distinct. The finite set of all non-repetitive paths ζ in S_k whose endpoints satisfy $\{\zeta_0,\zeta_{\ell(\zeta)}\}\subseteq\{x_1,\ldots,x_n\}$ will be denoted below by P_k. Suppose that for every path $\zeta\in\mathsf{P}_k$ we are given $c_1(\zeta),\ldots,c_{\ell(\zeta)}(\zeta)\in(0,\infty)$ such that for every $s,t\in\{0,\ldots,k\}$, $a\in\{1,\ldots,m_s\}$ and $b\in\{1,\ldots,m_t\}$ we have

$$\sum_{\zeta\in\mathsf{P}_k}\sum_{r=1}^{\ell(\zeta)}c_r(\zeta)\mathbf{1}_{\{(\zeta_{r-1},\zeta_r)=(x_a^s,x_b^t)\}}=E_{a,b}^{s,t}.$$

Then the inequality (86) can be rewritten as follows.

$$\sum_{\zeta\in\mathsf{P}_k}\sum_{r=1}^{\ell(\zeta)}c_r(\zeta)d_X(\zeta_{r-1},\zeta_r)^2\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}F_{i,j}d_X(x_i,x_j)^2. \tag{87}$$

By the triangle inequality and Cauchy-Schwarz, every $\zeta\in\mathsf{P}_k$ satisfies

$$d_X(\zeta_0,\zeta_{\ell(\zeta)})^2\leqslant\left(\sum_{r=1}^{\ell(\zeta)}\frac{1}{\sqrt{c_r(\zeta)}}\cdot\sqrt{c_r(\zeta)}d_X(\zeta_{r-1},\zeta_r)\right)^2\leqslant\left(\sum_{r=1}^{\ell(\zeta)}\frac{1}{c_r(\zeta)}\right)\sum_{r=1}^{\ell(\zeta)}c_r(\zeta)d_X(\zeta_{r-1},\zeta_r)^2. \tag{88}$$

By combining (87) and (88) we therefore see that

$$\sum_{\zeta\in\mathsf{P}_k}\frac{d_X(\zeta_0,\zeta_{\ell(\zeta)})^2}{\sum_{r=1}^{\ell(\zeta)}\frac{1}{c_r(\zeta)}}\leqslant\sum_{i=1}^{n}\sum_{j=1}^{n}F_{i,j}d_X(x_i,x_j)^2. \tag{89}$$

Recall that by the definition of P_k, the endpoints $\zeta_0,\zeta_{\ell(\zeta)}$ of any path $\zeta\in\mathsf{P}_k$ are in $\{x_1,\ldots,x_n\}$. It therefore follows from (89) that if we define for every $i,j\in\{1,\ldots,n\}$

$$G_{i,j}\stackrel{\mathrm{def}}{=}\sum_{\substack{\zeta\in\mathsf{P}_k\\(\zeta_0,\zeta_{\ell(\zeta)})=(x_i,x_j)}}\frac{1}{\sum_{r=1}^{\ell(\zeta)}\frac{1}{c_r(\zeta)}}, \tag{90}$$

then the following quadratic metric inequality, which generalizes (82), holds true in every Hadamard space (X,d_X).

$$\sum_{\substack{i,j\in\{1,\ldots,n\}\\G_{i,j}>F_{i,j}}}(G_{i,j}-F_{i,j})d_X(x_i,x_j)^2\leqslant\sum_{\substack{i,j\in\{1,\ldots,n\}\\F_{i,j}>G_{i,j}}}(F_{i,j}-G_{i,j})d_X(x_i,x_j)^2.$$

Question 34. Is it true that for every $D\in[1,\infty)$ there exists some $C(D)\in[1,\infty)$ such that a metric space (X,d_X) embeds with distortion at most $C(D)$ into some Hadamard space provided

$$\sum_{\substack{i,j\in\{1,\ldots,n\}\\G_{i,j}>F_{i,j}}}(G_{i,j}-F_{i,j})d_X(x_i,x_j)^2\leqslant D^2\sum_{\substack{i,j\in\{1,\ldots,n\}\\F_{i,j}>G_{i,j}}}(F_{i,j}-G_{i,j})d_X(x_i,x_j)^2,$$

for every $n\in\mathbb{N}$, every $x_1,\ldots,x_n\in X$ and every $\{F_{i,j},G_{i,j}\}_{i,j\in\{1,\ldots,n\}}$ as in (90)? Here we are considering all those $\{F_{i,j},G_{i,j}\}_{i,j\in\{1,\ldots,n\}}$ that are obtained from the construction that is described in

Section 5.2.1 and Section 5.2.2, i.e., ranging over all the possible choices of weights $\mu_{s,a}^{k+1,b}, \Gamma_{a,b}^{s,t}, c_r(\zeta)$ that were introduced in the course of this construction.

We conjecture that the answer to Question 34 is positive. It may even be the case that one could take $C(D) = D$ in Question 34. A negative answer here would be of great interest, since it would require finding a family of quadratic metric inequalities that does not follow (even up to a constant factor) from the above hierarchy of inequalities.

6. Remarks on Question 14

Focusing for concreteness on the case $p = 2$ of Question 14, recall that we are asking whether every n-point metric space (X, d_X) satisfies

$$c_{\mathcal{P}_2(\mathbb{R}^3)}(X) \lesssim \sqrt{\log n}. \tag{91}$$

The conclusion of Theorem 1, i.e., the fact that the 1/2-snowflake of every finite metric space embeds with $O(1)$ distortion into $\mathcal{P}_2(\mathbb{R}^3)$, does not on its own imply (91). Indeed, let $\sqrt{\ell_\infty}$ denote the 1/2-snowflake of ℓ_∞. Then the 1/2-snowflake of every finite metric space embeds isometrically into $\sqrt{\ell_\infty}$. However, it is standard to check that if for $n \in \mathbb{N}$ we let P_n denote the set $\{1, \ldots, n\} \subseteq \mathbb{R}$, equipped with the metric inherited from $\mathbb{R}$, then $c_{\sqrt{\ell_\infty}}(P_n) \gtrsim \sqrt{n}$. Thus, despite the fact that $\sqrt{\ell_\infty}$ is 1/2-snowflake universal, the distortion of n-point metric spaces in $\sqrt{\ell_\infty}$ can grow much faster than the rate of $\sqrt{\log n}$ that we desire in (91). Nevertheless, $\sqrt{\ell_\infty}$ is not an especially convincing example in our context, since it does not contain rectifiable curves (which is essentially the reason for the lower bound $c_{\sqrt{\ell_\infty}}(P_n) \gtrsim \sqrt{n}$), while $\mathcal{P}_2(\mathbb{R}^3)$ is an Alexandrov space of nonnegative curvature.

Note that $c_{\mathcal{P}_2(\mathbb{R}^3)}(X) \lesssim \log n$ for every n-point metric space (X, d_X), so $\mathcal{P}_2(\mathbb{R}^3)$ certainly does not exhibit the bad behavior that we described above for embeddings into $\sqrt{\ell_\infty}$. This logarithmic upper bound follows from the fact that $c_{\mathcal{P}_2([0,1])}(X) \lesssim \log n$, so in fact $c_{\mathcal{P}_p(Y)}(X) \lesssim \log n$ for every metric space (Y, d_Y) that contains a geodesic segment and every n-point metric space (X, d_X). The bound $c_{\mathcal{P}_2([0,1])}(X) \lesssim \log n$ is a consequence of Bourgain's embedding theorem [12] combined with the easy fact that every finite subset of ℓ_2 embeds with distortion 1 into $\mathcal{P}_2([0,1])$. To check the latter assertion, take any $X \subseteq \ell_2$ of cardinality n. We may assume without loss of generality that $X \subseteq \mathbb{R}^n$. Denoting

$$M \stackrel{\mathrm{def}}{=} 1 + \max_{x\in X} \max_{j\in\{1,\ldots,n-1\}} |x_{j+1} - x_j|,$$

define $f : X \to \mathcal{P}_2(\mathbb{R})$ by $f(x) \stackrel{\mathrm{def}}{=} \sum_{j=1}^n \delta_{x_j+Mj}$. The choice of M ensures that the sequence $\{x_j + Mj\}_{j=1}^n$ is strictly increasing, so for $x, y \in X$ the optimal transportation between $f(x)$ and $f(y)$ assigns the point mass at $x_j + Mj$ to the point mass at $y_j + Mj$ for every $j \in \{1, \ldots, n\}$. This shows that $\mathsf{W}_2(f(x), f(y)) = \|x - y\|_2$. Since all the measures $\{f(x)\}_{x\in X}$ are supported on a bounded interval, by rescaling we obtain a distortion 1 embedding of X into $\mathcal{P}_2([0,1])$.

An example that is more interesting in our context than $\sqrt{\ell_\infty}$, though still somewhat artificial, is the space $(\ell_2 \oplus \sqrt{\ell_\infty})_2$. This space is 1/2-snowflake universal (since it contains an isometric copy of $\sqrt{\ell_\infty}$) and also every n-point metric space (X, d_X) satisfies $c_{(\ell_2\oplus\sqrt{\ell_\infty})_2}(X) \lesssim \log n$ (by Bourgain's theorem [12], since $(\ell_2 \oplus \sqrt{\ell_\infty})_2$ contains an isometric copy of ℓ_2). However, we shall prove below the following lemma which shows that the conclusion of Question 14 fails for $(\ell_2 \oplus \sqrt{\ell_\infty})_2$.

Lemma 35. *For arbitrarily large $n \in \mathbb{N}$ there exists an n-point metric space (X_n, d_{X_n}) that satisfies*

$$c_{(\ell_2\oplus\sqrt{\ell_\infty})_2}(X_n) \gtrsim \log n.$$

Of course, $(\ell_2\oplus\sqrt{\ell_\infty})_2$ is still more pathological than $\mathcal{P}_2(\mathbb{R}^3)$ (in particular, not every pair of points in $(\ell_2\oplus\sqrt{\ell_\infty})_2$ can be joined by a rectifiable curve), and we lowered here the asymptotic growth rate of the largest possible distortion of an n-point metric space from the $O(\sqrt{n})$ of $\sqrt{\ell_\infty}$ to the $O(\log n)$ of $(\ell_2\oplus\sqrt{\ell_\infty})_2$ by artificially inserting a copy of ℓ_2. Nevertheless, the proof of Lemma 35 below illuminates the fact that in order to prove that Question 14 has a positive answer one would need to use properties of the Alexandrov space $\mathcal{P}_2(\mathbb{R}^3)$ that go beyond those that we isolated so far, and in particular it provides a concrete sequence of finite metric spaces for which the conclusion of Question 14 is at present unknown; see Question 36 below.

Before proving Lemma 35, we set some notation. For a finite connected graph $G=(V_G,E_G)$, the shortest-path metric that G induces on V_G is denoted by d_G. For $k\in\mathbb{N}$, denote the k-fold subdivision of G by $\Sigma_k(G)=(V_{\Sigma_k(G)},E_{\Sigma_k(G)})$, i.e., $\Sigma_k(G)$ is obtained from G by replacing each edge $e\in E_G$ by a path consisting of k edges joining the endpoints of e (the interiors of these paths are disjoint for distinct $e,e'\in E_G$). Thus $|V_{\Sigma_k(G)}|=|V_G|+(k-1)|E_G|$. Note that the metric induced on $V_G\subseteq V_{\Sigma_k(G)}$ by the shortest-path metric $d_{\Sigma_k(G)}$ of $\Sigma_k(G)$ is a rescaling of d_G by a factor of k, i.e.,

$$\forall\, x,y\in V_G\subseteq V_{\Sigma_k(G)},\qquad d_{\Sigma_k(G)}(x,y)=kd_G(x,y). \tag{92}$$

Suppose that G is d-regular for some $d\in\mathbb{N}$. The normalized adjacency matrix of G, i.e., the $V_G\times V_G$ matrix whose entry at $u,v\in V_G$ equals $1/d$ if $\{u,v\}\in E_G$ and equals 0 otherwise, is denoted A_G. The largest eigenvalue of the symmetric stochastic matrix A_G equals 1, and the second largest eigenvalue of A_G is denoted $\lambda_2(G)$.

Proof of Lemma 35. Fix $d,n\in\mathbb{N}$. We shall show that if $G=(V_G,E_G)$ is an n-vertex d-regular graph then

$$c_{(\ell_2\oplus\sqrt{\ell_\infty})_2}(\Sigma_k(G))\gtrsim\min\left\{\sqrt{\frac{k\log n}{\log d}},\sqrt{1-\lambda_2(G)}\cdot\frac{\log n}{\log d}\right\}. \tag{93}$$

In particular, for, say, $d=3$ and $\lambda_2(G)\leqslant 99/100$, if $k\asymp\log n$ then

$$c_{(\ell_2\oplus\sqrt{\ell_\infty})_2}(\Sigma_k(G))\asymp\log n\asymp\log|V_{\Sigma_k(G)}|.$$

This implies the validity of Lemma 35 because arbitrarily large graphs with the above requirements are well-known to exist (see e.g. [35]).

To prove (93), take $f:V_{\Sigma_k(G)}\to(\ell_2\oplus\sqrt{\ell_\infty})_2$ and suppose that there exist $s,D\in(0,\infty)$ such that for every $x,y\in V_{\Sigma_k(G)}$ we have

$$sd_{\Sigma_k(G)}(x,y)\leqslant d_{(\ell_2\oplus\sqrt{\ell_\infty})_2}(f(x),f(y))\leqslant Dsd_{\Sigma_k(G)}(x,y).$$

Our goal is to bound D from below. Writing $f(x)=(g(x),h(x))$ for every $x\in V_{\Sigma_k(G)}$, our assumption is that for every distinct $x,y\in V_{\Sigma_k(G)}$,

$$1\leqslant\frac{\|g(x)-g(y)\|_2^2+\|h(x)-h(y)\|_\infty}{s^2d_{\Sigma_k(G)}(x,y)^2}\leqslant D^2. \tag{94}$$

For $x,y\in V_{\Sigma_k(G)}$ with $\{x,y\}\in E_{\Sigma_k(G)}$ by (94) we have $\|h(x)-h(y)\|_\infty\leqslant s^2D^2=s^2D^2d_{\Sigma_k(G)}(x,y)$. Thus $h:V_{\Sigma_k(G)}\to\ell_\infty$ is s^2D^2-Lipschitz, and therefore

$$\frac{1}{n^2}\sum_{x,y\in V_G}\|h(x)-h(y)\|_\infty\leqslant\frac{s^2D^2}{n^2}\sum_{x,y\in V_G}d_{\Sigma_k(G)}(x,y)$$
$$\stackrel{(92)}{=}\frac{ks^2D^2}{n^2}\sum_{x,y\in V_G}d_G(x,y)\leqslant ks^2D^2\bigg(\frac{1}{n^2}\sum_{x,y\in V_G}d_G(x,y)^2\bigg)^{\frac12}. \tag{95}$$

Consequently, by (94) once more we have

$$\frac{1}{n^2}\sum_{x,y\in V_G}\|g(x)-g(y)\|_2^2 \overset{(94)}{\geqslant} \frac{s^2}{n^2}\sum_{x,y\in V_G} d_{\Sigma_k(G)}(x,y)^2 - \frac{1}{n^2}\sum_{x,y\in V_G}\|h(x)-h(y)\|_\infty$$
$$\overset{(95)}{\geqslant} \frac{k^2s^2}{n^2}\sum_{x,y\in V_G} d_G(x,y)^2 - ks^2D^2\bigg(\frac{1}{n^2}\sum_{x,y\in V_G} d_G(x,y)^2\bigg)^{\frac12}. \qquad (96)$$

At the same time, by the equivalent formulation of spectral gap in terms of a Poincaré inequality (see e.g. [29, Section 9.1] or [53, 61]),

$$\frac{1}{n^2}\sum_{x,y\in V_G}\|g(x)-g(y)\|_2^2 \leqslant \frac{1}{1-\lambda_2(G)}\cdot\frac{2}{|E_G|}\sum_{\substack{x,y\in V_G\\ \{x,y\}\in E_G}}\|g(x)-g(y)\|_2^2$$
$$\overset{(94)}{\leqslant} \frac{s^2D^2}{1-\lambda_2(G)}\cdot\frac{2}{|E_G|}\sum_{\substack{x,y\in V_G\\ \{x,y\}\in E_G}} d_{\Sigma_k(G)}(x,y)^2 \overset{(92)}{=} \frac{2s^2k^2D^2}{1-\lambda_2(G)}. \qquad (97)$$

By contrasting (96) with (97) we deduce that

$$D \gtrsim \min\left\{\bigg(\frac{1-\lambda_2(G)}{n^2}\sum_{x,y\in V_G} d_G(x,y)^2\bigg)^{\frac12}, \bigg(\frac{k^2}{n^2}\sum_{x,y\in V_G} d_G(x,y)^2\bigg)^{\frac14}\right\}.$$

This lower bound on D implies the desired estimate (93) since by a standard (and simple) counting argument (see e.g. [53, page 193]) the fact that G has n vertices and is d-regular implies that

$$\bigg(\frac{1}{n^2}\sum_{x,y\in V_G} d_G(x,y)^2\bigg)^{\frac12} \gtrsim \frac{\log n}{\log d}. \qquad \square$$

Question 36. Suppose that G is an n-vertex 3-regular graph with $\lambda_2(G)\leqslant 99/100$. What is the asymptotic growth rate of

$$c_{\mathcal{P}_2(\mathbb{R}^3)}\big(\Sigma_{\lceil \log n\rceil}(G)\big)?$$

At present, the best known upper bound on this quantity is $O(\log n)$, while Question 14 predicts that it is $O(\sqrt{\log n})$. Obtaining any $o(\log n)$ upper bound would be interesting here.

Acknowledgments. A. A. was supported in part by the NSF and the Simons Foundation. A. N. was supported in part by the BSF, the Packard Foundation and the Simons Foundation. O. N. was supported in part by the ISF and the European Union's Seventh Framework Programme. Some of the results of the present article were announced in the extended abstract [2] that was presented in the 43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016). We refer to [2] for algorithmic context and applications of our results to theoretical computer science.

References

[1] S. Alexander, V. Kapovitch, and A. Petrunin. Alexandrov meets Kirszbraun. In *Proceedings of the Gökova Geometry-Topology Conference 2010*, pages 88–109. Int. Press, Somerville, MA, 2011.

[2] A. Andoni, A. Naor, and O. Neiman. Impossibility of Sketching of the 3D Transportation Metric with Quadratic Cost. In Y. R. Ioannis Chatzigiannakis, Michael Mitzenmacher and D. Sangiorgi, editors, *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 83:1–83:14, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[3] T. Austin and A. Naor. On the bi-Lipschitz structure of Wasserstein spaces. Preprint, 2015.

[4] T. Austin, A. Naor, and Y. Peres. The wreath product of $\mathbb{Z}$ with $\mathbb{Z}$ has Hilbert compression exponent $\frac{2}{3}$. *Proc. Amer. Math. Soc.*, 137(1):85–90, 2009.

[5] K. Ball. Markov chains, Riesz transforms and Lipschitz maps. *Geom. Funct. Anal.*, 2(2):137–172, 1992.

[6] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science (Burlington, VT, 1996)*, pages 184–193. IEEE Comput. Soc. Press, Los Alamitos, CA, 1996.

[7] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On metric Ramsey-type phenomena. *Ann. of Math. (2)*, 162(2):643–709, 2005.

[8] I. D. Berg and I. G. Nikolaev. On a distance characterization of A. D. Aleksandrov spaces of nonpositive curvature. *Dokl. Akad. Nauk*, 414(1):10–12, 2007.

[9] I. D. Berg and I. G. Nikolaev. Quasilinearization and curvature of Aleksandrov spaces. *Geom. Dedicata*, 133:195–218, 2008.

[10] A. Beurling and L. Ahlfors. The boundary correspondence under quasiconformal mappings. *Acta Math.*, 96:125–142, 1956.

[11] J. Bourgain. New Banach space properties of the disc algebra and H^∞. *Acta Math.*, 152(1-2):1–48, 1984.

[12] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.

[13] J. Bourgain. The metrical interpretation of superreflexivity in Banach spaces. *Israel J. Math.*, 56(2):222–230, 1986.

[14] J. Bourgain and V. Milman. Dichotomie du cotype pour les espaces invariants. *C. R. Acad. Sci. Paris Sér. I Math.*, 300(9):263–266, 1985.

[15] M. R. Bridson and A. Haefliger. *Metric spaces of non-positive curvature*, volume 319 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[16] J. Briët, A. Naor, and O. Regev. Locally decodable codes and the failure of cotype for projective tensor products. *Electron. Res. Announc. Math. Sci.*, 19:120–130, 2012.

[17] S. M. Buckley, K. Falk, and D. J. Wraith. Ptolemaic spaces and CAT(0). *Glasg. Math. J.*, 51(2):301–314, 2009.

[18] Y. Burago, M. Gromov, and G. Perel′man. A. D. Aleksandrov spaces with curvatures bounded below. *Uspekhi Mat. Nauk*, 47(2(284)):3–51, 222, 1992.

[19] H. Busemann. *The geometry of geodesics*. Academic Press Inc., New York, N. Y., 1955.

[20] F. Critchley and B. Fichet. The partial order by inclusion of the principal classes of dissimilarity on a finite set, and some of their basic properties. In *Classification and dissimilarity analysis*, volume 93 of *Lecture Notes in Statist.*, pages 5–65. Springer, New York, 1994.

[21] G. David and S. Semmes. *Fractured fractals and broken dreams*, volume 7 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press, Oxford University Press, New York, 1997. Self-similar geometry through metric and measure.

[22] J. Ding, J. R. Lee, and Y. Peres. Markov type and threshold embeddings. *Geom. Funct. Anal.*, 23(4):1207–1229, 2013.

[23] P. Enflo. On the nonexistence of uniform homeomorphisms between L_p-spaces. *Ark. Mat.*, 8:103–105, 1969.

[24] P. Enflo. Uniform homeomorphisms between Banach spaces. In *Séminaire Maurey-Schwartz (1975–1976), Espaces, L^p, applications radonifiantes et géométrie des espaces de Banach, Exp. No. 18*, page 7. Centre Math., École Polytech., Palaiseau, 1976.

[25] T. Foertsch, A. Lytchak, and V. Schroeder. Nonpositive curvature and the Ptolemy inequality. *Int. Math. Res. Not. IMRN*, (22):Art. ID rnm100, 15, 2007.

[26] T. Foertsch and V. Schroeder. Hyperbolicity, CAT(-1)-spaces and the Ptolemy inequality. *Math. Ann.*, 350(2):339–356, 2011.

[27] D. J. H. Garling. *Inequalities: a journey into linear analysis*. Cambridge University Press, Cambridge, 2007.

[28] G. Godefroy and N. J. Kalton. Lipschitz-free Banach spaces. *Studia Math.*, 159(1):121–141, 2003. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday.

[29] M. Gromov. Filling Riemannian manifolds. *J. Differential Geom.*, 18(1):1–147, 1983.

[30] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[31] M. Gromov. *Metric structures for Riemannian and non-Riemannian spaces*, volume 152 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1999. Based on the 1981 French original [MR0682063 (85e:53051)], With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[32] M. Gromov. CAT(κ)-spaces: construction and concentration. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 280(Geom. i Topol. 7):100–140, 299–300, 2001.

[33] M. Gromov. Random walk in random groups. *Geom. Funct. Anal.*, 13(1):73–146, 2003.

[34] S. Har-Peled and M. Mendel. Fast construction of nets in low-dimensional metrics and their applications. *SIAM J. Comput.*, 35(5):1148–1184 (electronic), 2006.

[35] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006.

[36] J. Jost. *Nonpositive curvature: geometric and analytic aspects*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997.

[37] D. C. Kay. The ptolemaic inequality in Hilbert geometries. *Pacific J. Math.*, 21:293–301, 1967.

[38] S. Khot and A. Naor. Nonembeddability theorems via Fourier analysis. *Math. Ann.*, 334(4):821–852, 2006.

[39] B. Kleiner and B. Leeb. Rigidity of quasi-isometries for symmetric spaces and Euclidean buildings. *Inst. Hautes Études Sci. Publ. Math.*, (86):115–197 (1998), 1997.

[40] S. Kwapień and A. Pełczyński. Absolutely summing operators and translation-invariant spaces of functions on compact abelian groups. *Math. Nachr.*, 94:303–340, 1980.

[41] T. J. Laakso. Plane with A_∞-weighted metric not bi-Lipschitz embeddable to $\mathbb{R}^N$. *Bull. London Math. Soc.*, 34(6):667–676, 2002.

[42] V. Lafforgue. Un renforcement de la propriété (T). *Duke Math. J.*, 143(3):559–602, 2008.

[43] V. Lafforgue. Propriété (T) renforcée banachique et transformation de Fourier rapide. *J. Topol. Anal.*, 1(3):191–206, 2009.

[44] U. Lang, B. Pavlović, and V. Schroeder. Extensions of Lipschitz maps into Hadamard spaces. *Geom. Funct. Anal.*, 10(6):1527–1553, 2000.

[45] U. Lang and C. Plaut. Bilipschitz embeddings of metric spaces into space forms. *Geom. Dedicata*, 87(1-3):285–307, 2001.

[46] U. Lang and V. Schroeder. Kirszbraun's theorem and metric spaces of bounded curvature. *Geom. Funct. Anal.*, 7(3):535–560, 1997.

[47] N. Lebedeva and A. Petrunin. Curvature bounded below: a definition a la Berg-Nikolaev. *Electron. Res. Announc. Math. Sci.*, 17:122–124, 2010.

[48] J. R. Lee, M. Mendel, and A. Naor. Metric structures in L_1: dimension, snowflakes, and average distortion. *European J. Combin.*, 26(8):1180–1190, 2005.

[49] J. R. Lee and A. Naor. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geom. Funct. Anal.*, 14(4):745–747, 2004.

[50] B. Liao. Strong Banach property (T) for simple algebraic groups of higher rank. *J. Topol. Anal.*, 6(1):75–105, 2014.

[51] N. Linial, A. Magen, and A. Naor. Girth and Euclidean distortion. *Geom. Funct. Anal.*, 12(2):380–394, 2002.

[52] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[53] J. Matoušek. On embedding expanders into l_p spaces. *Israel J. Math.*, 102:189–197, 1997.

[54] J. Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[55] B. Maurey. Type, cotype and K-convexity. In *Handbook of the geometry of Banach spaces, Vol. 2*, pages 1299–1332. North-Holland, Amsterdam, 2003.

[56] M. Mendel. Metric dichotomies. In *Limits of graphs in group theory and computer science*, pages 59–76. EPFL Press, Lausanne, 2009.

[57] M. Mendel and A. Naor. Metric cotype. *Ann. of Math. (2)*, 168(1):247–298, 2008.

[58] M. Mendel and A. Naor. Maximum gradient embeddings and monotone clustering. *Combinatorica*, 30(5):581–615, 2010.

[59] M. Mendel and A. Naor. A note on dichotomies for metric transforms. Available at `http://arxiv.org/abs/1102.1800`, 2011.

[60] M. Mendel and A. Naor. Markov convexity and local rigidity of distorted metrics. *J. Eur. Math. Soc. (JEMS)*, 15(1):287–337, 2013.

[61] M. Mendel and A. Naor. Nonlinear spectral calculus and super-expanders. *Publ. Math. Inst. Hautes Études Sci.*, 119:1–95, 2014.

[62] M. Mendel and A. Naor. Expanders with respect to Hadamard spaces and random graphs. *Duke Math. J.*, 164(8):1471–1548, 2015.

[63] M. Mimura. Sphere Equivalence, Banach Expanders, and Extrapolation. *Int. Math. Res. Not. IMRN*, (12):4372–4391, 2015.

[64] A. Naor. A phase transition phenomenon between the isometric and isomorphic extension problems for Hölder functions between L_p spaces. *Mathematika*, 48(1-2):253–271 (2003), 2001.

[65] A. Naor. An application of metric cotype to quasisymmetric embeddings. In *Metric and differential geometry*, volume 297 of *Progr. Math.*, pages 175–178. Birkhäuser/Springer, Basel, 2012.
[66] A. Naor. An introduction to the Ribe program. *Jpn. J. Math.*, 7(2):167–233, 2012.
[67] A. Naor. Comparison of metric spectral gaps. *Anal. Geom. Metr. Spaces*, 2:1–52, 2014.
[68] A. Naor and Y. Peres. Embeddings of discrete groups and the speed of random walks. *Int. Math. Res. Not. IMRN*, pages Art. ID rnn 076, 34, 2008.
[69] A. Naor and Y. Peres. L_p compression, traveling salesmen, and stable walks. *Duke Math. J.*, 157(1):53–108, 2011.
[70] A. Naor, Y. Peres, O. Schramm, and S. Sheffield. Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces. *Duke Math. J.*, 134(1):165–197, 2006.
[71] A. Naor and G. Schechtman. Remarks on non linear type and Pisier's inequality. *J. Reine Angew. Math.*, 552:213–236, 2002.
[72] A. Naor and G. Schechtman. Planar earthmover is not in L_1. *SIAM J. Comput.*, 37(3):804–826 (electronic), 2007.
[73] A. Naor and L. Silberman. Poincaré inequalities, embeddings, and wild groups. *Compos. Math.*, 147(5):1546–1572, 2011.
[74] I. Newman and Y. Rabinovich. A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.*, 29(1):77–81, 2003.
[75] P. W. Nowak. Poincaré inequalities and rigidity for actions on Banach spaces. *J. Eur. Math. Soc. (JEMS)*, 17(3):689–709, 2015.
[76] S.-I. Ohta. Markov type of Alexandrov spaces of non-negative curvature. *Mathematika*, 55(1-2):177–189, 2009.
[77] S.-i. Ohta and M. Pichot. A note on Markov type constants. *Arch. Math. (Basel)*, 92(1):80–88, 2009.
[78] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
[79] N. Ozawa. A note on non-amenability of $\mathcal{B}(l_p)$ for $p = 1, 2$. *Internat. J. Math.*, 15(6):557–565, 2004.
[80] A. Pełczyński. *Banach spaces of analytic functions and absolutely summing operators*. American Mathematical Society, Providence, R.I., 1977. Expository lectures from the CBMS Regional Conference held at Kent State University, Kent, Ohio, July 11–16, 1976, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 30.
[81] G. Pisier. Ensembles de Sidon et espaces de cotype 2. In *Séminaire sur la Géométrie des Espaces de Banach (1977–1978)*, pages Exp. No. 14, 12. École Polytech., Palaiseau, 1978.
[82] G. Pisier. Factorization of operator valued analytic functions. *Adv. Math.*, 93(1):61–125, 1992.
[83] S. T. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I*. Probability and its Applications (New York). Springer-Verlag, New York, 1998. Theory.
[84] Y. G. Reshetnyak. Inextensible mappings in a space of curvature no greater than k. *Sibirsk. Math. J.*, 9:683–689, 1968.
[85] T. Sato. An alternative proof of Berg and Nikolaev's characterization of CAT(0)-spaces via quadrilateral inequality. *Arch. Math. (Basel)*, 93(5):487–490, 2009.
[86] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.
[87] K. T. Sturm. Metric spaces of lower bounded curvature. *Exposition. Math.*, 17(1):35–47, 1999.
[88] K.-T. Sturm. Probability measures on metric spaces of nonpositive curvature. In *Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002)*, volume 338 of *Contemp. Math.*, pages 357–390. Amer. Math. Soc., Providence, RI, 2003.
[89] K.-T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.
[90] D. Thurston. Length inequalities in trees and CAT(0) spaces. MathOverflow. URL:http://mathoverflow.net/q/163706 (version: 2014-04-22).
[91] N. Tomczak-Jaegermann. The moduli of smoothness and convexity and the Rademacher averages of trace classes $S_p(1 \leqslant p < \infty)$. *Studia Math.*, 50:163–182, 1974.
[92] P. Tukia and J. Väisälä. Quasisymmetric embeddings of metric spaces. *Ann. Acad. Sci. Fenn. Ser. A I Math.*, 5(1):97–114, 1980.
[93] N. Varopoulos. Une remarque sur les ensembles de Helson. *Duke Math. J.*, 43(2):387–390, 1976.
[94] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
[95] J. H. Wells and L. R. Williams. *Embeddings and extensions in analysis*. Springer-Verlag, New York, 1975. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 84.
[96] P. Wojtaszczyk. *Banach spaces for analysts*, volume 25 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1991.

[97] T. Yokota. A rigidity theorem in Alexandrov spaces with lower curvature bound. *Math. Ann.*, 353(2):305–331, 2012.

Computer Science Department, Columbia University, 500 West 120th Street, Room 450, MC0401, New York, NY 10027, USA.

E-mail address: andoni@cs.columbia.edu

Mathematics Department, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA.

E-mail address: naor@math.princeton.edu

Department of Computer Science, Ben-Gurion University of the Negev, P.O.B 653 Be'er Sheva 84105, Israel

E-mail address: neimano@cs.bgu.ac.il