An optimal algorithm for average distance in typical regular graphs*

Alexandros Eskenazis† Manor Mendel‡ Assaf Naor§

Abstract. We design a deterministic algorithm that, given n points in a *typical* constant degree regular graph, queries $O(n)$ distances to output a constant factor approximation to the average distance among those points, thus answering a question posed in [52]. Our algorithm uses the method of [52] to construct a sequence of constant degree graphs that are expanders with respect to certain nonpositively curved metric spaces, together with a new rigidity theorem for metric transforms of nonpositively curved metric spaces. The fact that our algorithm works for typical (uniformly random) constant degree regular graphs rather than for all constant degree graphs is unavoidable, thanks to the following impossibility result that we obtain: For every fixed $k \in \mathbb{N}$, the approximation factor of any algorithm for average distance that works for all constant degree graphs and queries $o(n^{1+1/k})$ distances must necessarily be at least $2(k+1)$. This matches the upper bound attained by the algorithm that was designed for general finite metric spaces in [9]. Thus, any algorithm for average distance in constant degree graphs whose approximation guarantee is less than 4 must query $\Omega(n^2)$ distances, any such algorithm whose approximation guarantee is less than 6 must query $\Omega(n^{3/2})$ distances, any such algorithm whose approximation guarantee less than 8 must query $\Omega(n^{4/3})$ distances, and so forth, and furthermore there exist algorithms achieving those parameters.

1 Introduction

Indyk systematically investigated [37] fast algorithms for geometric computations involving distances. One of the main problems that was featured in [37] is approximating the average distance[1]

$$\frac{1}{n^2}\sum_{(i,j)\in[n]^2} d_{\mathcal{M}}(x_i,x_j) \tag{1.1}$$

among n points $x_1,\ldots,x_n \in \mathcal{M}$ in a given metric space $(\mathcal{M},d_{\mathcal{M}})$. As there are n^2 summands in (1.1), such an algorithm is considered sublinear if it performs $o(n^2)$ distance queries. Obtaining a finite approximation factor entails making at least $n-1$ distance queries since otherwise the graph whose edges are the pairs of points that were queried would be disconnected, so by varying the distances between its connected components one can obtain two markedly different metric spaces that the algorithm cannot distinguish.

Here, we will study deterministic algorithms for the question stated above. In combination with the mathematics that we will develop for that purpose (belonging to metric embeddings and Alexandrov geometry), this investigation will lead to the resolution of a range of questions that were left open in the literature. In order to explain these matters, it is beneficial to first ground the discussion (and set notation/terminology) by recalling the following formal definition of the computational model that will be treated herein. Its generality entails that the lower bound that we will obtain is quite a strong statement, while the algorithm that we will design belongs to the least expressive aspect of this framework, namely, it will be non-adaptive and, in fact, it will be a "universal approximator" per terminology from [9] that we will soon recall.

Definition 1.1. *Let $\mathcal{F}$ be a family of metric spaces. For $n,m \in \mathbb{N}$ and $\alpha \geq 1$, a deterministic α-approximation algorithm for the average distance in $\mathcal{F}$ that makes m distance queries on inputs of size n consists of functions*

$$\Big\{\mathsf{Pair}_i : \Big(\tbinom{[n]}{2}\times[0,\infty)\Big)^{i-1} \to \tbinom{[n]}{2}\Big\}_{i=1}^{m} \quad \text{and} \quad \mathsf{Estimate} : \Big(\tbinom{[n]}{2}\times[0,\infty)\Big)^{m} \to [0,\infty), \tag{1.2}$$

*Complete proofs of the results that are covered in the present extended abstract, together with further results and applications to pure mathematics, as well as additional context, appear in [27].

†CNRS, Inst. de Math. de Jussieu, Sorbonne U., France (alexandros.eskenazis@imj-prg.fr, https://www.alexandroseskenazis.com).
‡Dep. of Math. and Comp. Sci., The Open U. of Israel, Israel (manorme@openu.ac.il, https://sites.google.com/site/mendelma).
§Math. Dep., Princeton U., Princeton, NJ, USA (naor@math.princeton.edu, https://web.math.princeton.edu/~naor/).

[1] Throughout the ensuing text, we will use the common notations $[n]=\{1,\ldots,n\}$ and $\binom{[n]}{2}=\{\{a,b\}\subseteq[n]:\ a\neq b\}$ for $n\in\mathbb{N}$.

where Pair_1 *is understood to be a fixed element* $\mathsf{Pair}_1 = \{a_1, b_1\} \in \binom{[n]}{2}$ *that satisfies the following property.*

Given a metric space $(\mathcal{M}, d_{\mathcal{M}}) \in \mathcal{F}$ *and* $x_1, \ldots, x_n \in \mathcal{M}$*, define* $\{a_1, b_1\}, \{a_2, b_2\}, \ldots, \{a_m, b_m\} \in \binom{[n]}{2}$ *inductively (recalling that* $\{a_1, b_1\}$ *has already been fixed above) by setting, for every* $i \in [m-1]$*,*

$$\{a_i, b_i\} \stackrel{\mathrm{def}}{=} \mathsf{Pair}_i \Big(\big(\{a_1, b_1\}, d_{\mathcal{M}}(x_{a_1}, x_{b_1})\big), \big(\{a_2, b_2\}, d_{\mathcal{M}}(x_{a_2}, x_{b_2})\big), \ldots, \big(\{a_{i-1}, b_{i-1}\}, d_{\mathcal{M}}(x_{a_{i-1}}, x_{b_{i-1}})\big) \Big). \tag{1.3}$$

We then require that[2]

$$1 \le \frac{\mathsf{Estimate}\Big(\big(\{a_1, b_1\}, d_{\mathcal{M}}(x_{a_1}, x_{b_1})\big), \ldots, \big(\{a_m, b_m\}, d_{\mathcal{M}}(x_{a_m}, x_{b_m})\big) \Big)}{\frac{1}{n^2} \sum_{(i,j) \in [n]^2} d_{\mathcal{M}}(x_i, x_j)} \le \alpha. \tag{1.4}$$

A non-adaptive *deterministic* α*-approximation algorithm for average distance in* $\mathcal{F}$ *which makes* m *distance queries on size* n *inputs, is the restrictive case of the above setup in which* $\mathsf{Pair}_1, \ldots, \mathsf{Pair}_m$ *are constant functions, so their images are, respectively,* m *pairs of indices* $\{a_1, b_1\}, \ldots, \{a_m, b_m\} \in \binom{[n]}{2}$*. The output is then a function of the* $d_{\mathcal{M}}$*-distances between the pairs* $\{x_{a_1}, x_{b_1}\}, \ldots, \{x_{a_m}, x_{b_m}\} \subseteq \{x_1, \ldots, x_n\}$ *of points whose indices were decided upfront and not adapted to the specific input* $x_1, \ldots, x_n \in \mathcal{M}$*.*

When we discuss α*-approximation algorithms for average distance that make* m *distance queries on inputs of size* n *without specifying the underlying family* $\mathcal{F}$*, we will tacitly mean that* $\mathcal{F}$ *consists of all metric spaces; as any finite metric space is isometric to a subset of* ℓ_∞*, we may take, in this case,* $\mathcal{F} = \{\ell_\infty\}$*.*

The standard interpretation of (1.3) is that the algorithm performs the following "approximation game" in m rounds. Starting by querying the $d_{\mathcal{M}}$-distance between x_{a_1} and x_{b_1}, based only on the answer it receives, the algorithm computes a new pair of points x_{a_2}, x_{b_2} and queries the $d_{\mathcal{M}}$-distance between them. In round $i+1$, the algorithm selects a new pair of points $x_{a_{i+1}}, x_{b_{i+1}}$ and queries their $d_{\mathcal{M}}$ distance, where that pair is a function of only the (ordered) sequence of pairs queried by the algorithm and their $d_{\mathcal{M}}$ distances (the responses to the queries) in the preceding i rounds. After m rounds, the algorithm outputs an estimate for the average distance, which is required to be a function of only the (ordered) sequence of the pairs it queried and the responses to those queries throughout this procedure.

Although the new algorithmic results obtained herein only concern deterministic algorithms, we wish to mention randomized algorithms when referring to previous results in the literature. Thus, in the context of Theorem 1.1, a randomized α-approximation algorithm for the average distance in $\mathcal{F}$ that makes m distance queries on inputs of size n is the same as in Theorem 1.1, except that the functions in (1.2) have an additional variable ω from some probability space $(\Omega, \mathbb{P})$, and it is required that the approximation guarantee (1.4) holds with probability at least, say, 2/3. In the non-adaptive special case, $\mathsf{Pair}_1, \ldots, \mathsf{Pair}_m$ depend only on ω. In this context, $\mathcal{F}$ is understood to be all metric spaces if it is not specified.

Indyk considered the simple randomized non-adaptive algorithm that is obtained by averaging over a uniformly random sampling of $O(n/\varepsilon^{7/2})$ pairs for some $0 < \varepsilon < 1$. He proved that this straightforward procedure yields a $1+\varepsilon$ factor approximation with constant probability. Barhum, Goldreich, and Shraibman [9] improved Indyk's analysis to $O(n/\varepsilon^2)$ queries. Goldreich and Ron [31, 32] studied a restricted version of the above problem in which one wishes to approximate the average distance among *all* vertices of a given *unweighted connected* graph on n vertices, showing that averaging over a uniformly random sampling of $O(\sqrt{n}/\varepsilon^2)$ pairs of vertices yields a $(1+\varepsilon)$-approximation algorithm with constant probability.[3]

Deterministic algorithms for average distance estimation were broached in [9], where an especially simple example of non-adaptive algorithms, called *universal approximators*, was studied. Per Definition 6 in [9], given $\alpha \ge 1$, an α-universal approximator in a family $\mathcal{F}$ of metric spaces of size m for inputs of size n consists of m unordered pairs of indices $\{a_1, b_1\}, \ldots, \{a_m, b_m\} \in \binom{[n]}{2}$ and a scaling factor $\sigma > 0$ such that for every metric space $(\mathcal{M}, d_{\mathcal{M}}) \in \mathcal{F}$ and every $x_1, \ldots, x_n \in \mathcal{M}$ the output of the algorithm is

$$\mathsf{Estimate}_{\mathrm{BGS}}\big(d_{\mathcal{M}}(x_{a_1}, x_{b_1}), \ldots, d_{\mathcal{M}}(x_{a_m}, x_{b_m})\big) \stackrel{\mathrm{def}}{=} \sigma \sum_{\ell=1}^{m} d_{\mathcal{M}}(x_{a_\ell}, x_{b_\ell}),$$

[2]If the denominator in (1.4) vanishes (i.e., when $x_1 = \ldots = x_n$), then we require that the numerator in (1.4) also vanishes.

[3]This algorithm depends crucially on estimating the average over all pairs of vertices and the graph being unweighted, whereas in our model there is $\Omega(n)$ lower bound even for adaptive randomized algorithms.

and (1.4) holds. That is, $n^{-2}\sum_{(i,j)\in[n]^2} d_{\mathcal{M}}(x_i,x_j) \le \sigma\sum_{\ell=1}^{m} d_{\mathcal{M}}(x_{a_\ell},x_{b_\ell}) \le \alpha n^{-2}\sum_{(i,j)\in[n]^2} d_{\mathcal{M}}(x_i,x_j)$ for every metric space $(\mathcal{M},d_{\mathcal{M}})\in\mathcal{F}$ and every $x_1,\ldots,x_n\in\mathcal{M}$. As before, if $\mathcal{F}$ is not mentioned, then it will be assumed tacitly to be all the possible metric spaces. The following theorem was proved in [9]:

Theorem 1.2. *For $k,n\in\mathbb{N}$ with $k\ge 2$ there is an n-vertex $(2k)$-universal approximator of size $O(kn^{1+1/k})$.*

In particular, Theorem 1.2 shows that for every fixed integer $k\ge 2$ there is a (nonadaptive) deterministic $(2k)$-approximation algorithm for the average distance that makes $O(n^{1+1/k})$ distance queries on inputs of size n. Theorem 1.3 below, which is one of the main results of the present work, provides a matching impossibility statement. It should be noted that, in addition to proving Theorem 1.2, a (non-matching) lower bound on the size of universal approximators was proved in [9]. Ruling out any algorithm whatsoever (in the aforementioned oracle model) rather than only universal approximators is conceptually an entirely different matter. In fact, the concrete adversarial metric / widget that is used in [9] fails to fool even some non-adaptive algorithms, while our proof of Theorem 1.3 builds an implicit adversarial metric (algorithm-dependent) as a solution to an auxiliary optimization problem. An overview of the proof of Theorem 1.3 appears in section 2 below, and its complete details are presented in [27].

Theorem 1.3. *Fix $k\in\mathbb{N}$ and $\alpha\ge 1$. If for all $n\in\mathbb{N}$, there is a deterministic α-approximation algorithm for the average distance that makes $o(n^{1+1/k})$ distance queries on inputs of size n, then necessarily $\alpha\ge 2(k+1)$.*

By combining Theorem 1.2 and Theorem 1.3, we obtain the following description of the approximation landscape for the average distance. Any algorithm for the average distance with an approximation guarantee less than 4 must query $\Omega(n^2)$ distances; any such algorithm with an approximation guarantee less than 6 must query $\Omega(n^{3/2})$ distances; any such algorithm whose approximation guarantee is less than 8 must query $\Omega(n^{4/3})$ distances, etc. Furthermore, there are algorithms that attain these parameters. See Figure 1.1 for a graphical representation of the lower bound versus the upper bound.

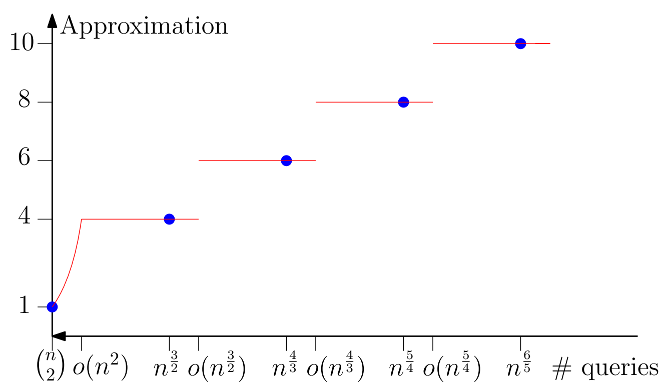

Figure 1.1: The blue dots correspond to the available algorithms per Theorem 1.2. The red lines correspond to the impossibility statement of Theorem 1.3. Notice that the graph's horizontal axis decreases as one moves from left to right, and it introduces a macroscopic gap between $O(n^{1+1/k})$ queries and $o(n^{1+1/k})$ queries.

Remark 1.4. Similar-looking (but not identical!) tight approximation-to-size trade-offs are postulated as famous conjectures on related problems, such as spanners [2] and approximate distance oracles [73]. However, the tight trade-off that we obtain here is an unconditional theorem because, in contrast to the aforementioned examples, our lower bound does not assume the validity of the long-standing Erdős girth conjecture [24]. In our opinion, this aspect of Theorem 1.3 should be further explored.

Thanks to Theorem 1.3, we now know that there does not exist a deterministic constant approximation algorithm for the average distance that makes $O(n)$ distance queries on inputs of size n. By the simple fact contained in Theorem 1.5 below, this impossibility result implies the same statement for constant approximation algorithms for the average distance in $\mathsf{Reg}(d)$, which makes $O(n)$ distance queries on inputs of size n. Here for each $d\in\{3,4,\ldots,\}$ we let $\mathsf{Reg}(d)$ denote the family of metric spaces that consists of all finite connected d regular graphs, equipped with their shortest-path metric. A detailed (straightforward) proof of Theorem 1.5 appears in [27].

PROPOSITION 1.5. *For every $\varepsilon > 0$ and every $d \in \{3, 4, \dots\}$, any finite metric space embeds with distortion $1+\varepsilon$ into some finite connected d-regular graph, equipped with its shortest-path metric.*

The (standard; see, e.g., [47, Chapter 15]) terminology that is used in Theorem 1.5 (as well as throughout the ensuing text) is that an embedding $f : \mathcal{M} \to \mathcal{N}$ of a metric space $(\mathcal{M}, d_{\mathcal{M}})$ into a metric space $(\mathcal{N}, d_{\mathcal{N}})$ has (bi-Lipschitz) distortion at most $D \geq 1$ if there exists (a scaling factor) $s > 0$ such that

$$\forall x, y \in \mathcal{M}, \qquad d_{\mathcal{M}}(x, y) \leq s d_{\mathcal{N}}\big(f(x), f(y)\big) \leq D d_{\mathcal{M}}(x, y). \tag{1.5}$$

The $\mathcal{N}$-distortion of $\mathcal{M}$ is a numerical invariant that is commonly denoted $\mathsf{c}_{\mathcal{N}}(\mathcal{M})$, following [45], and is defined as the infimum over those $D \geq 1$ such that there exists an embedding $f : \mathcal{M} \to \mathcal{N}$ whose distortion is at most D. Sometimes, one also needs to use the notations $\mathsf{c}_{(\mathcal{N},d_{\mathcal{N}})}(\mathcal{M}, d_{\mathcal{M}})$ or $\mathsf{c}_{\mathcal{N}}(\mathcal{M}, d_{\mathcal{M}})$ instead of $\mathsf{c}_{\mathcal{N}}(\mathcal{M})$ when the corresponding underlying metrics are not clear from the context. Such situations in which multiple metrics on the same space must be considered, and will indeed occur below. When $\mathcal{M}$ is finite and $p \geq 1$, it is common to use the shorthand $\mathsf{c}_p(\mathcal{M}) = \mathsf{c}_{\ell_p}(\mathcal{M})$.

It is possible to bypass the aforementioned hardness by considering deterministic approximation algorithms for the average distance that work when the input is allowed to be an arbitrary n-tuple of vertices in a *typical* finite connected regular graph rather than in any such graph. The precise formulation of this phenomenon appears in the following theorem, which is another main result of the present work:

THEOREM 1.6. *For $N \in \mathbb{N}$ and $d \in \{3, 4, \dots, \}$, let $\mathcal{G}_{N,d}$ be the set of all the (isomorphism classes of) N-vertex d-regular graphs. There exists $\mathcal{F}_{N,d} \subseteq \mathcal{G}_{N,d}$ that consists of connected graphs (equipped with their shortest path metric) that satisfy $\lim_{N\to\infty} |\mathcal{F}_{N,d}|/|\mathcal{G}_{N,d}| = 1$, such that the family $\mathcal{F}_d = \bigcup_{N=1}^{\infty} \mathcal{F}_{N,d}$ of metric spaces has the following property. There exists a deterministic algorithm that takes as input $n \in \mathbb{N}$ and outputs in $O(n)$ time m pairs $\{a_1, b_1\}, \dots, \{a_m, b_m\} \subseteq [n]$, where $m = O(n)$, such that*[4]

$$\forall \mathsf{G} = (\mathsf{V_G}, \mathsf{E_G}) \in \mathcal{F}_d,\ \forall x_1, \dots, x_n \in \mathsf{V_G}, \qquad \frac{1}{m}\sum_{\ell=1}^{m} d_{\mathsf{G}}(x_{a_\ell}, x_{b_\ell}) \asymp \frac{1}{n^2}\sum_{(i,j)\in[n]^2} d_{\mathsf{G}}(x_i, x_j), \tag{1.6}$$

where in (1.6)*, and throughout what follows, d_{G} denotes the shortest-path metric on a connected graph G. In particular, there exists a non-adaptive deterministic $O(1)$-approximation algorithm for the average distance in $\mathcal{F}_d$ which makes $O(n)$ distance queries on inputs of size n (in contrast to Theorem 1.3).*

The precursor to Theorem 1.6 is [52, Theorem 1.2] or [54, Theorem 2.1], providing the same statement as Theorem 1.6 except that the distances appearing in (1.6) are *squared*. The obvious question if Theorem 1.6 holds as stated above for approximating the actual average distance (in accordance with the prior literature on this topic) rather than approximating the average quadratic distance was a central problem that was left open in [52, 54]. Now, it is resolved (positively) by Theorem 1.6.

Remark 1.7. Even though the impossibility of a constant approximation with $m = O(n)$ queries that Theorem 1.3 establishes was not known at the time of writing of [52, 54], in hindsight we see that [52, 54] proposed a different way to cope with the nonexistence of deterministic algorithms. The aforementioned sublinear average distance approximation problem falls into the category of multiple known algorithmic tasks (in an oracle model) for which constant factor approximation is provably impossible for deterministic algorithms (with suitable limitations on the number of oracle queries), yet it is achievable using randomized algorithms. A notable instance of this type of situation is the problem of approximating the volume of a convex body in $\mathbb{R}^n$ that is given by a weak membership oracle (see [36] for the relevant terminology), for which of Bárány and Füredi established [8] the impossibility of constant-factor approximation in oracle-polynomial time, yet Dyer, Frieze and Kannan proved [23] that such a randomized algorithm does exist (a curiosity of this analogy is that the chronology is reversed for average distance approximation, namely, in that setting a randomized algorithm was known to exist [37] before hardness for deterministic algorithms was established). The idea of [52, 54], which is taken up by Theorem 1.6,

[4]We will use the following (standard) conventions for asymptotic notation, in addition to the usual $O(\cdot), \Omega(\cdot), \Theta(\cdot)$ notation. Given $a, b > 0$, by writing $a \lesssim b$ or $b \gtrsim a$ we mean $a \leq \kappa b$ for a universal constant $\kappa > 0$, and $a \asymp b$ stands for $(a \lesssim b) \wedge (b \lesssim a)$.

is that when it is known that there is no deterministic algorithm, instead of asking for a randomized algorithm that succeeds (with high probability) on all possible inputs, one could still demand that the algorithm will be deterministic but require it to efficiently provide a good approximation on *every* sub-structure of a typical input. In the present setting "typical" is understood to be uniformly random and "sub-structure" is an arbitrary subset, but one could conceive of other versions of such phenomena to investigate (e.g. typical inputs could arise from various probabilistic models that are thought to mimic real-world instances). For example, in the above setting of volume computation, one could consider for some integers $1 \le n \le N$ a random convex body $K \subseteq \mathbb{R}^N$ (a well-studied example of this type is Gluskin bodies [29] and variants thereof [72], which are convex hulls of randomly selected points in $\mathbb{R}^N$, but there are many more such possibilities in the literature), and then demand that a deterministic algorithm yields a good approximation for the n-dimensional volume of *every* n-dimensional section of K. It could be a worthwhile research direction to study in multiple settings such deterministic algorithmic tasks that are required to succeed on every sub-structure of a random input.

Passing from squared distances to the distances themselves is not merely a technical matter that was left unresolved in the literature; indeed, it was speculated in [52, 54] (specifically, see Remark 2.3 there) that achieving this would require a substantially new idea beyond those of [52, 54], because the proof in [52, 54] of the aforementioned quadratic variant makes heavy use of a metric space-valued martingale inequality[5] which fails to hold (even when the underlying metric space is an interval in the real line) if the distances are raised to power 1 (or any power $0 < p < 2$) rather than power 2.

The above prediction materializes herein through the geometric rigidity result in Theorem 1.8 below. It provides an ingredient that makes it possible to apply the methodology of [52, 54] to obtain Theorem 1.6. Although the present extended abstract focuses on the algorithmic context, it should be noted that Theorem 1.8 is of geometric interest in its own right and implies solutions to long-standing questions in metric geometry. Those will be briefly mentioned in Theorem 1.10 below and thoroughly treated in [27]. Before stating Theorem 1.8, we will next recall a modicum of (standard) background.

We will use throughout the common terminology (see, e.g., the monographs [13, 75, 21, 16]) that a function $\varphi : [0,\infty) \to [0,\infty)$ is called a *metric transform* if φ is nondecreasing, concave and satisfies $\varphi(0) = 0$. This ensures that $(\mathcal{M}, \varphi \circ d_{\mathcal{M}})$ is a metric space for every metric space $(\mathcal{M}, d_{\mathcal{M}})$.

A metric space $(\mathcal{M}, d_{\mathcal{M}})$ is said to be nonpositively curved in the sense of Alexandrov if it is a geodesic metric space, i.e., for any two points x and y in $\mathcal{M}$ there exists a constant speed geodesic $\gamma : [0,1] \to \mathcal{M}$ joining x to y (namely, $d_{\mathcal{M}}(\gamma(t), x) = t d_{\mathcal{M}}(x,y)$ and $d_{\mathcal{M}}(\gamma(t), y) = (1-t) d_{\mathcal{M}}(x,y)$ for every $0 \le t \le 1$), and the distances between points along geodesic triangles in $\mathcal{M}$ are bounded from above by the corresponding distances in a triangle with the same edge lengths that resides in the Euclidean plane $\mathbb{R}^2$. This is equivalent to requiring that for every $x, y, z \in \mathcal{M}$, if $\gamma : [0,1] \to \mathcal{M}$ is a geodesic of constant-speed that joins x to y, then

$$\forall t \in [0,1], \qquad d_{\mathcal{M}}(\gamma(t), z)^2 \le t d_{\mathcal{M}}(x,z)^2 + (1-t) d_{\mathcal{M}}(y,z)^2 - t(1-t) d_{\mathcal{M}}(x,y)^2. \tag{1.7}$$

Thorough treatments of this important, useful and extensively studied notion appears in, e.g., [39, 15, 18, 71, 35, 10]. Following [33], metric spaces that are nonpositively curved in the sense of Alexandrov are also called CAT(0) spaces (shorthand for "Cartan, Alexandrov and Toponogov"), and complete CAT(0) spaces are commonly called Hadamard spaces.

We can now state Theorem 1.8, which is the main geometric contribution of the present work; an outline of its proof is contained in section 3 below, and all details appear in [27]. It should be clarified at the outset that even though our proof of Theorem 1.6 relies on Theorem 1.8, the relevance of Theorem 1.8 to Theorem 1.6 is roundabout, and readers who are exposed for the first time to this material should not expect it to be obvious how Theorem 1.8 could lead to Theorem 1.6. The complete details of this deduction appear in [27], and we will indicate some of its ingredients below, but the entirety of the proof is beyond the scope of the present extended abstract.

THEOREM 1.8. *Let $\varphi : [0,\infty) \to [0,\infty)$ be a metric transform, and let $(\mathcal{M}, d_{\mathcal{M}})$ be a* CAT(0) *metric space. Then, the metric space $(\mathcal{M}, \varphi \circ d_{\mathcal{M}})$ is embedded with constant distortion in some (other)* CAT(0) *space.*

Our proof of Theorem 1.6 follows the approach developed in [52, 54] to obtain its aforementioned quadratic variant. Theorem 1.8 provides the previously missing ingredient that enables us to implement the ideas of [52, 54]

[5] A version due to [51] of Pisier's classical martingale cotype inequality [66] for (suitably defined) *nonlinear* martingales.

for this purpose, ideas that, on their own, failed to produce Theorem 1.6. We do not see how to use for this purpose the results of [52, 54] as a "black box," so we need to adjust multiple details within the *proofs of* [52, 54] in order to be able to combine them with the new geometric input provided by Theorem 1.8 so as to derive Theorem 1.6.

The starting point of the reasoning is the important realization of [9] that universal approximators are related to *expander graphs*; indeed, using this connection, [9] obtained for any $0 < \varepsilon < 1$ a deterministic non-adaptive $(1+\varepsilon)$-approximation algorithm for the average distance in the family $\mathcal{F}$ of subsets of Hilbert space that makes $O(n/\varepsilon^2)$ queries on inputs of size n. The link between expanders and universal approximators is the point of view from which [52, 54] approached the construction of a constant approximation algorithm of size $O(n)$ for the average squared distance in subsets of random regular graphs, with the "twist" that now one encounters the need to consider this question from the perspective of the theory of nonlinear spectral gaps and expanders with respect to metric spaces; this is an extensively studied topic that is recalled in [27] (information on this rich mathematical research direction can be found in [48, 74, 64, 38, 43, 67, 61, 51, 53, 57, 54, 55, 58, 26, 20, 59, 25, 60], and applications of it to algorithm design can be found in [52, 5]). The special expanders that were used in the construction of [52, 54] are variants of the zigzag construction of Reingold, Vadhan and Wigderson [68], relying on its adaptation to the metric space setting in [53, 51].

The method of [52, 54] relies on tracking how nonlinear spectral gaps with respect to certain metric spaces evolve under a suitable iteration of the zigzag product; this utilized inequalities for nonlinear martingales that are available by [51] on CAT(0) spaces (among others). For that reason, even though the ultimate goal is to obtain a nonlinear spectral gap with respect to (discrete) random regular graphs, in [52, 54] those graphs were considered as subsets of a larger (continuous) CAT(0) space on which the aforementioned martingale analysis can be performed. Theorem 1.8 comes to our aid at this point by taking the metric transform $\varphi(t) = \sqrt{t}$ of the aforementioned larger auxiliary metric space, which Theorem 1.8 embeds into an even larger CAT(0) on which one could hope to apply the martingale tools developed in [53, 51, 52, 54]. However, this observation itself is insufficient and a better understanding of the connection between metric transforms of random regular graphs and CAT(0) spaces is needed; the complete details of this appear in [27].

Question 1.9. Our proof of Theorem 1.6 uses methods from analysis and geometry (in addition to graph theory and probability), even though its statement is solely about elementary combinatorics. It remains an intriguing challenge to find a combinatorial proof of Theorem 1.6 that does not make such a circuitous excursion to mathematical areas that are (seemingly) distant from the setting of Theorem 1.6. In addition to its intrinsic interest, this goal is likely to be relevant to related questions that remain open; an example of such a question (posed by J. Kleinberg) is discussed in [54, § 2].

Remark 1.10. Even though the present extended abstract is devoted to algorithmic matters, it would be remiss of us not to mention that in [27] we derive other applications of Theorem 1.8 to pure mathematics. For example, by combining Theorem 1.8 with [34, 41, 4], we prove that there exists a metric space $(\mathcal{M}, d_{\mathcal{M}})$ that embeds with constant distortion into a metric space $(\mathcal{X}, d_{\mathcal{X}})$ of nonpositive Alexandrov curvature, and also $(\mathcal{M}, d_{\mathcal{M}})$ embeds with constant distortion into a metric space $(\mathcal{Y}, d_{\mathcal{Y}})$ of nonnegative Alexandrov curvature[6], yet $(\mathcal{M}, d_{\mathcal{M}})$ does not admit a coarse embedding into a Hilbert space (see [69, 35, 62] for background on the well-studied notion of coarse embeddings, though in [27] we obtain an even stronger impossibility result). This is in contrast to a classical result of Wilson [76] and Blumenthal [12] (see also [13, pages 122–128] or [46, Section 7]) which implies that a geodesic metric space that embeds isometrically into both a metric space of nonpositive Alexandrov curvature and a metric space of nonnegative Alexandrov curvature must be isometric to a subset of a Hilbert space. As another example of an application of Theorem 1.8, in [27] we resolve (negatively) a question that Ding, Lee and Peres posed in [22, Question 1.13] by proving that there exists a metric space which has Markov type 2 (per [6]), *sharp* metric cotype 2 (per [50]), and is Markov 2-convex (per [44]), yet, it does not embed coarsely into a Hilbert space. This rules out (in a strong form) an approach (in accordance with the Ribe program [14, 40, 56, 7, 63, 30, 58]) toward a possible metric version of Kwapień's theorem [42] that was investigated in [22]. As a third example of an application of Theorem 1.8, in [27] we address an old folklore question in the theory of metric embeddings about the possible ℓ_1 distortion growth rates, by proving that for any metric transform φ there exists a Hadamard space $\mathcal{M}_\varphi$ for which $\sup\{\mathsf{c}_1(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}_\varphi \wedge |\mathcal{C}| \le n\}$ is bounded from above and from below by positive universal

[6]Nonnegative Alexandrov curvature for a geodesic metric space corresponds to reversing the inequality in (1.7).

constant multiples of $\varphi(\log(n+1))/\varphi(1)$. For embeddings into a Hilbert space, we are currently able to obtain this result under the additional assumption that $\varphi(t) \gtrsim \varphi(1)\sqrt{\log t}$. In particular, we prove that for every $0 \le \theta \le 1$ there exists a Hadamard space $\mathcal{M}_\theta$ for which $\sup\{\mathsf{c}_2(\mathcal{C}) : \mathcal{C} \subseteq \mathcal{M}_\theta \ \wedge |\mathcal{C}| \le n\}$ is bounded from above and from below by positive universal constant multiples of $(\log(n+1))^\theta$; this statement was not previously known when $\theta \notin \{0, \frac{1}{2}, 1\}$, even if one relaxes it by allowing $\mathcal{M}_\theta$ to be any geodesic metric space.

2 Overview of the proof of Theorem 1.3 Due to their length, the complete details of the proof of Theorem 1.3—our (optimal) lower bounds for deterministic approximation algorithms for average distance that work for all metric spaces—are deferred to [27]. This section presents an outline of this proof.

The following theorem implies Theorem 1.3 in a straightforward manner (see below):

THEOREM 2.1. *Fix $k \in \mathbb{N}$. Continuing with the notation of Theorem 1.1, suppose that there is a deterministic α-approximation algorithm for the average distance which makes $m = o(n^{1+1/k})$ distance queries on inputs of size n. Thus, we are given functions as in* (1.2) *which satisfy the properties described in Theorem 1.1 when $\mathcal{F}$ is the family of all possible metric spaces. Then, there exist*

$$\overline{d}, \underline{d} : [n]^2 \to [0, k+1] \qquad \text{and} \qquad \{a_1, b_1\}, \ldots, \{a_m, b_m\} \in \tbinom{[n]}{2}, \tag{2.1}$$

with the following properties:

(I) Both $\overline{d}$ and $\underline{d}$ are metrics on $[n]$ (of diameter at most $k+1$);

(II) For every $i \in [m]$ we have $\overline{d}(a_i, b_i) = \underline{d}(a_i, b_i) \stackrel{\mathrm{def}}{=} w(a_i, b_i)$;

(III) The recursive relation (1.3) *holds for $(x_1, \ldots, x_n) = (1, \ldots, n)$, i.e.,*

$$\forall i \in [m-1], \qquad \{a_{i+1}, b_{i+1}\} = \mathsf{Pair}_{i+1}\Big(\big(\{a_1, b_1\}, w(a_1, b_1)\big), \ldots, \big(\{a_i, b_i\}, w(a_i, b_i)\big)\Big); \tag{2.2}$$

(IV) There exists $X \subseteq [n]^2$ with $|X| = o(n^2)$ such that $\overline{d}(x, y) = k+1$ for every $(x, y) \in [n]^2 \smallsetminus X$;

(V) There exists $Y \subseteq [n]^2$ with $|Y| = o(n^2)$ such that $\underline{d}(x, y) = \frac{1}{2}$ for every $(x, y) \in [n]^2 \smallsetminus Y$ with $x \ne y$.

We check that Theorem 2.1 indeed implies Theorem 1.3, by contrasting item (I), item (II) and item (III) in the conclusion of Theorem 2.1 with Theorem 1.1 we see that

$$\begin{aligned}\frac{1}{n^2}\sum_{(i,j)\in[n]^2} \overline{d}(i,j) &\stackrel{(1.4)}{\le} \mathsf{Estimate}\Big(\big(\{a_1, b_1\}, \overline{d}(x_{a_1}, x_{b_1})\big), \ldots, \big(\{a_m, b_m\}, \overline{d}(a_m, b_m)\big)\Big) \\ &= \mathsf{Estimate}\Big(\big(\{a_1, b_1\}, \underline{d}(x_{a_1}, x_{b_1})\big), \ldots, \big(\{a_m, b_m\}, \underline{d}(a_m, b_m)\big)\Big) \stackrel{(1.4)}{\le} \frac{\alpha}{n^2}\sum_{(i,j)\in[n]^2} \underline{d}(i,j).\end{aligned} \tag{2.3}$$

At the same time, by item (IV) we have

$$\frac{1}{n^2}\sum_{(i,j)\in[n]^2} \overline{d}(i,j) \ge \frac{n^2 - |X|}{n^2}(k+1) = \big(1 - o(1)\big)(k+1), \tag{2.4}$$

and by item (IV) we have (using the fact that the diameter of the metric $\underline{d}$ is at most $k+1$),

$$\frac{1}{n^2}\sum_{(i,j)\in[n]^2} \underline{d}(i,j) \le \frac{n^2 - |Y|}{n^2} \cdot \frac{1}{2} + \frac{|Y|}{n^2}\operatorname{diam}\big([n], \underline{d}\big) \le \frac{1 - o(1)}{2} + o(1)(k+1). \tag{2.5}$$

By contrasting (2.3) with (2.4) and (2.5) and letting $n \to \infty$, we conclude that $\alpha \ge 2(k+1)$, as required.

The remainder of this section will sketch the proof of Theorem 2.1. For that and throughout what follows, given an edge-weighted graph $\mathsf{G} = \big(\mathsf{V}, \mathsf{E}, w : \mathsf{E} \to [0, \infty)\big)$ we will denote by $d_{\mathsf{G}} : \mathsf{V}^2 \to [0, \infty]$ its associated shortest-path (extended) metric, with the convention that $d_{\mathsf{G}}(x, y) = \infty$ if $x, y \in \mathsf{V}$ belong to distinct connected components of G. The degree in G of a vertex $x \in \mathsf{V}$ will be denoted $\deg_{\mathsf{G}}(x) \in \mathbb{N} \cup \{0\}$.

We will define $\{a_1,b_1\},\ldots,\{a_i,b_i\}\in\binom{[n]}{2}$ and a metric d_i on $[n]$ by induction on $i\in[m]$. Per Theorem 1.1, the initial pair $\{a_1,b_1\}$ is fixed as part of the algorithm's parameters. Set $d_1(a_1,b_1)=1$. Assume inductively that $\{a_1,b_1\},\ldots,\{a_i,b_i\}$ and d_i have already been defined for $i\in[m-1]$.

To go from i to $i+1$ in our recursive definition, let the next pair $\{a_{i+1},b_{i+1}\}\in\binom{[n]}{2}$ be given by

$$\{a_{i+1},b_{i+1}\}\stackrel{\mathrm{def}}{=}\mathsf{Pair}_{i+1}\Big(\big(\{a_1,b_1\},d_i(a_1,b_1)\big),\ldots,\big(\{a_i,b_i\},d_i(a_i,b_i)\big)\Big). \tag{2.6}$$

Write $\mathsf{E}_i\stackrel{\mathrm{def}}{=}\big\{\{a_1,b_1\},\ldots,\{a_i,b_i\}\big\}$. Let $\mathsf{G}_i\stackrel{\mathrm{def}}{=}\big([n],\mathsf{E}_i,d_i\big)$ be the edge-weighted graph whose vertex set is $[n]$, whose edge set is E_i, and in which the weight of each edge $\{a,b\}\in\mathsf{E}_i$ equals $d_i(a,b)$. Next, define

$$\forall x\in[n],\qquad h_i(x)\stackrel{\mathrm{def}}{=}\min\left\{\left\lfloor\frac{k}{\log n}\log\Big(1+\frac{1}{\sqrt{m}}n^{\frac{k+1}{2k}}\deg_{\mathsf{G}_i}(x)\Big)\right\rfloor,k-1\right\}\in\{0,\ldots,k-1\}. \tag{2.7}$$

We now define $d_{i+1}(a,b)$ as follows for every $a,b\in[n]$ such that $\{a,b\}\in\big\{\{a_1,b_1\},\ldots,\{a_{i+1},b_{i+1}\}\big\}\stackrel{\mathrm{def}}{=}\mathsf{E}_{i+1}$:

$$d_{i+1}(a,b)\stackrel{\mathrm{def}}{=}\max\Big\{\min\big\{\max\{h_i(a),h_i(b)\}+1,d_{\mathsf{G}_i}(a,b)\big\},\max_{\{u,v\}\in\mathsf{E}_i}\big(d_i(u,v)-d_{\mathsf{G}_i}(u,a)-d_{\mathsf{G}_i}(v,b)\big)\Big\}. \tag{2.8}$$

Observe that $d_{i+1}(a,b)=d_{i+1}(b,a)$, so we may consider the weighted graph $\mathsf{G}_{i+1}\stackrel{\mathrm{def}}{=}\big([n],\mathsf{E}_{i+1},d_{i+1}\big)$. An elementary induction that is carried out in [27] shows that

$$\forall j\in[i+1],\qquad d_{i+1}(a_j,b_j)=d_{\mathsf{G}_{i+1}}(a_j,b_j)=d_{\mathsf{G}_j}(a_j,b_j)\le k. \tag{2.9}$$

Although $d_{i+1}(a,b)$ was initially defined in (2.8) only for those $(a,b)\in[n]^2$ such that $\{a,b\}\in\mathsf{E}_{i+1}$, by (2.9) the following definition provides an extension of d_{i+1} to all of $[n]^2$ which is also a metric on $[n]$:

$$\forall(a,b)\in[n]^2,\qquad d_{i+1}(a,b)\stackrel{\mathrm{def}}{=}\min\big\{d_{\mathsf{G}_{i+1}}(a,b),k+1\big\}. \tag{2.10}$$

This completes the inductive step from i to $i+1$.

Continuing with the proof of Theorem 2.1, we will take the metric $\overline{d}$ in (2.1) to be $\overline{d}\stackrel{\mathrm{def}}{=}d_m$. That is, it is obtained as the result of the inductive construction described above in step $i=m$. It indeed takes values in $[0,k+1]$, as required in (2.1), thanks to (2.10). The pairs $\{a_1,b_1\},\ldots,\{a_m,b_m\}$ in (2.1) are also given in step $i=m$.

Writing $w(a_i,b_i)\stackrel{\mathrm{def}}{=}\overline{d}(a_i,b_i)$ for $i\in[m]$ according to item (II) of Theorem 2.1, we see from (2.9) that $d_i(a_i,b_i)=d_{i+1}(a_i,b_i)=\ldots=d_m(a_i,b_i)=w(a_i,b_i)$ for every $i\in[m]$. Consequently, the desired equality (2.2) in item (III) of Theorem 2.1 coincides with (2.6).

For the ensuing discussion, it will be convenient to write $\mathsf{E}\stackrel{\mathrm{def}}{=}\mathsf{E}_m$ and $\mathsf{G}\stackrel{\mathrm{def}}{=}\mathsf{G}_m$. Thus, we have

$$\forall x,y\in[n],\qquad \overline{d}(x,y)=\min\big\{d_{\mathsf{G}}(x,y),k+1\big\}\in\{0,\ldots,k+1\}. \tag{2.11}$$

We will also use the notation $h\stackrel{\mathrm{def}}{=}h_m$, and let $B_{\mathsf{G}}(x,r)\stackrel{\mathrm{def}}{=}\{y\in[n]:\ d_{\mathsf{G}}(x,y)\le r\}$ be the closed ball with respect to the shortest-path metric d_{G} on the weighted graph G centered at $x\in[n]$ of radius $r\ge 0$. The following lemma on the geometry of G is proved in [27]:

LEMMA 2.2. *The growth rate of the size of* G*-balls of radius at most* $k-1$ *satisfies the following:*

$$\forall(x,r)\in[n]\times[k-1],\qquad |B_{\mathsf{G}}(x,r)|\lesssim\frac{\sqrt{m}}{n^{\frac{k-2r+1}{2k}}}. \tag{2.12}$$

Also, the following relation holds between the size of any G*-ball of radius* k *and the* G*-degree of its center:*

$$\forall x\in[n],\qquad \deg_{\mathsf{G}}(x)<\frac{\sqrt{m}}{n^{\frac{k+1}{2k}}}(n-1)\implies|B_{\mathsf{G}}(x,k)|\lesssim n^{\frac{1}{2}-\frac{1}{2k}}\sqrt{m}. \tag{2.13}$$

Item (IV) of Theorem 2.1 can be quickly deduced from the second part (2.13) of Theorem 2.2 for the following choice of $X \subseteq [n]^2$:

$$X \stackrel{\mathrm{def}}{=} \left\{(x,y) \in [n]^2 : \ \overline{d}(x,y) \le k\right\} \stackrel{(2.11)}{=} \left\{(x,y) \in [n]^2 : \ y \in B_{\mathsf{G}}(x,k)\right\}. \tag{2.14}$$

As $\overline{d}$ is integer-valued, the first equality in (2.14) gives that if $(x,y) \in [n]^2 \smallsetminus X$, then $\overline{d}(x,y) \ge k+1$. Next, set

$$H \stackrel{\mathrm{def}}{=} \left\{x \in [n] : \ \deg_{\mathsf{G}}(x) \ge \frac{\sqrt{m}}{n^{\frac{k+1}{2k}}}(n-1)\right\} \stackrel{(2.13)}{\Longrightarrow} \forall x \in [n] \smallsetminus H, \qquad |B_{\mathsf{G}}(x,k)| \lesssim n^{\frac{1}{2}-\frac{1}{2k}}\sqrt{m}. \tag{2.15}$$

Assuming that $n \ge 2$, we then have

$$m = |\mathsf{E}| = \frac{1}{2}\sum_{x\in[n]} \deg_{\mathsf{G}}(x) \ge \frac{\sqrt{m}}{2n^{\frac{k+1}{2k}}}(n-1)|H| \ge \frac{n^{\frac{k-1}{2k}}\sqrt{m}}{4}|H| \implies |H| \le \frac{4\sqrt{m}}{n^{\frac{k-1}{2k}}}. \tag{2.16}$$

$$|X| \stackrel{(2.14)}{\le} \sum_{x\in[n]} |B_{\mathsf{G}}(x,k)| \le |H|n + n\max_{x\in[n]\smallsetminus H} |B_{\mathsf{G}}(x,k)| \stackrel{(2.15)\wedge(2.16)}{\lesssim} n^{\frac{1}{2}+\frac{1}{2k}}\sqrt{m} + n^{\frac{3}{2}-\frac{1}{2k}}\sqrt{m} \asymp n^{\frac{3}{2}-\frac{1}{2k}}\sqrt{m}, \tag{2.17}$$

where the second step of (2.17) uses the trivial estimate $|B_{\mathsf{G}}(x,k)| \le n$ for $x \in H$ and the last step of (2.17) holds since $k \ge 1$. Due to the assumption $m = o(n^{1+1/k})$ of Theorem 2.1, it follows from (2.17) that $|X| = o(n^2)$. This concludes the verification of item (IV) of Theorem 2.1.

The second metric $\underline{d}$ in (2.1) is implicitly defined as any minimizer of the (linear) objective function

$$\left(d \in [0,\infty)^{[n]^2}\right) \mapsto \sum_{(x,y)\in[n]^2} d(x,y), \tag{2.18}$$

subject to the following system of (linear) constraints:

$$\begin{cases} \forall x \in [n], & d(x,x) = 0, \\ \forall x,y \in [n], & d(x,y) = d(y,x), \\ \forall x,y,z \in [n], & d(x,z) \le d(x,y) + d(y,z), \\ \forall \{x,y\} \in \mathsf{E}, & d(x,y) = w(x,y), \\ \forall x,y \in [n], & \min\left\{\overline{d}(x,y), \max\{h(x),h(y)\} + \frac{1}{2}\right\} \le d(x,y) \le \overline{d}(x,y). \end{cases} \tag{2.19}$$

The set of all those $d : [n]^2 \to [0,\infty)$ that satisfy (2.19) is nonempty since $\overline{d}$ belongs to it. Due to the last constraint in (2.19), this set is compact, so there is a minimizer $\underline{d}$ of (2.18). The first three constraints in (2.19) ensure that $\underline{d}$ is a metric on $[n]$. Since by (2.11) we know that $\overline{d}$ takes values in $[0,k+1]$, thanks to the last constraint in (2.19) we also know that $\underline{d}$ takes values in $[0,k+1]$, as required in (2.1). The fourth constraint in (2.19) ensures that item (II) of Theorem 2.1 holds. Hence, the proof Theorem 2.1 will be complete if we prove that its conclusion item (V) is true.

Remark 2.3. A naïve way to obtain $\underline{d}$ would be to consider a minimizer of (2.18) subject to the first four constraints in (2.19), as this would produce a (pseudo)metric[7] for which item (II) of Theorem 2.1 holds, and it has the smallest possible average distance among all such (pseudo)metrics. However, it turns out that adding the fifth constraint in (2.19) yields further information that facilitates the subsequent analysis.

Below, closed balls that are induced by the metrics $\overline{d}$ and $\underline{d}$ are denoted

$$\forall (x,r) \in [n]\times\mathbb{R}, \qquad B_{\overline{d}}(x,r) \stackrel{\mathrm{def}}{=} \left\{y \in [n] : \ \overline{d}(x,y) \le r\right\} \qquad \text{and} \qquad B_{\underline{d}}(x,r) \stackrel{\mathrm{def}}{=} \left\{y \in [n] : \ \underline{d}(x,y) \le r\right\}.$$

[7]The remaining requirement $d(x,y) \ge 0$ that is needed for d to be a pseudometric follows by taking $x = z$ in the third constraint in (2.19) while applying the first constraint in (2.19).

It turns out that the following subset of $[n]^2$ satisfies Conclusion ((V)) of Theorem 2.1:

$$Y \stackrel{\mathrm{def}}{=} \big(U\times[n]\big)\cup\big([n]\times U\big)\cup W\cup \widetilde{W}\cup\big\{(x,y)\in[n]^2:\ \{x,y\}\in\mathsf{E}\big\}\cup\big\{(x,x):\ x\in[n]\big\}, \tag{2.20}$$

where $U\subseteq[n]$ and $W,\widetilde{W}\subseteq[n]^2$ in (2.20) are defined as follows:

$$\begin{aligned} &U \stackrel{\mathrm{def}}{=} \bigcup_{x\in[n]} B_{\underline{d}}\big(x,h(x)-1\big),\\ &W \stackrel{\mathrm{def}}{=} \bigcup_{u\in[n]} \Big(B_{\underline{d}}\big(u,h(u)\big)\times\big\{v\in[n]:\ \{u,v\}\in\mathsf{E}\ \text{ and }\ w(u,v)=h(u)+1\big\}\Big),\\ &\widetilde{W} \stackrel{\mathrm{def}}{=} \bigcup_{u\in[n]} \Big(\big\{v\in[n]:\ \{u,v\}\in\mathsf{E}\ \text{ and }\ w(u,v)=h(u)+1\big\}\times B_{\underline{d}}\big(u,h(u)\big)\Big)=\big\{(x,y)\in[n]^2:\ (y,x)\in W\big\}. \end{aligned} \tag{2.21}$$

To prove that Y satisfies the requirements of item (V) of Theorem 2.1, one must show that $|Y|=o(n^2)$ and that $\underline{d}(x,y)=\frac12$ for distinct $x,y\in[n]$ such that $(x,y)\in[n]^2\smallsetminus Y$. The latter is the content of Theorem 2.4 below, whose detailed proof appears in [27]. This is a key step in which the choice of $\underline{d}$ as a minimizer of (2.18) is used extensively through variational reasoning, as discussed below.

Lemma 2.4. *If $x,y\in[n]^2$ and $(x,y)\notin Y$, then $\underline{d}(x,y)\in\big\{0,\frac12\big\}$.*

The proof of Theorem 2.4 is by contradiction. It starts by taking the distinct $(x,y)\in[n]^2\smallsetminus Y$ with $\underline{d}(x,y)\neq 1/2$, and furthermore $\underline{d}(x,y)$ is *maximal* among all those $x,y\in[n]$ with this property. Observe that the first inequality in the fifth constraint in (2.19) ensures that $\underline{d}(u,v)\ge 1/2$ for every distinct $u,v\in[n]$, so we necessarily have $\underline{d}(x,y)>1/2$. Recalling the definition (2.20) of Y, the fact that $(x,y)\notin Y$ implies $x,y\notin U$. By the definition of U in (2.21), this implies $h(x)=h(y)=0$, since otherwise if, say, $h(x)\ge 1$ (recall that $h\ge 0$ is integer valued), then we have $x\in B_{\underline{d}}(x,h(x)-1)\subseteq U$. Hence, the first inequality in the fifth constraint in (2.19) is strict. Furthermore, since $Y\supseteq\{(a,b)\in[n]^2:\ \{a,b\}\in\mathsf{E}\}$ by (2.20), we know that $\{x,y\}\notin\mathsf{E}$, so the fourth constraint in (2.19) does not apply to $\underline{d}(x,y)$. Because $\underline{d}$ is a minimizer of (2.18) subject to the system of constraints (2.19), the aforementioned considerations imply that there must exist $z\in[n]\smallsetminus\{x,y\}$ such that $\underline{d}(x,z)=\underline{d}(x,y)+\underline{d}(y,z)$, because otherwise it would have been possible to reduce $\underline{d}(x,y)$ without changing the rest of the values of $\underline{d}$ while ensuring that all of the constraints in (2.19) are still satisfied (this is where the strictness of the lower bound on $\underline{d}(x,y)$ that appears in the fifth constraint in (2.19) is used), which contradicts the fact that $\underline{d}$ is a minimizer of (2.18) subject to those constraints.

The above process is iterated in [27] to obtain $(u,v)\in[n]^2$ which are endpoints of a discrete geodesic with respect to $\underline{d}$ that contains x,y (thus $\underline{d}(u,x)+\underline{d}(x,y)+\underline{d}(y,v)=\underline{d}(u,v)$), and is maximal with respect to inclusion. As before, one of the lower bounds on $\underline{d}$ that appear in the system of constraints (2.19) must hold as equality, but this time, due to the maximality of the aforementioned geodesic whose endpoints are u and v, it cannot be the third constraint in (2.19), i.e., the triangle inequality. Hence, either the first inequality in the fifth constraint in (2.19) holds as equality, or $\{u,v\}\in\mathsf{E}$, i.e., the fourth constraint in (2.19) applies. From here, a case analysis that is performed in [27] leads to the desired contradiction.

To conclude the sketch of the proof of Theorem 2.1, it remains to explain why $|Y|=o(n^2)$. Observe that

$$\forall x,y\in[n],\qquad \underline{d}(x,y)<\max\{h(x),h(y)\}+\tfrac12 \implies \underline{d}(x,y)=\overline{d}(x,y). \tag{2.22}$$

Indeed, (2.22) is implied by the fifth constraint in (2.19). Now,

$$\forall x\in[n],\forall r\in[h(x)],\qquad B_{\underline{d}}(x,r)\stackrel{(2.22)}{=}B_{\overline{d}}(x,r)\stackrel{(2.7)\wedge(2.11)}{=}B_{\mathsf{G}}(x,r), \tag{2.23}$$

where the last step of (2.23) holds since $h(x)\le k-1$ for every $x\in[n]$ by (2.7) and $\overline{d}$ is given by (2.11). The a priori upper bound $h(x)\le k-1$ also allows us to apply the first part (2.12) of Theorem 2.2 to deduce using (2.23) that

$$\forall x\in[n],\forall r\in[h(x)],\qquad \big|B_{\underline{d}}(x,r)\big|\lesssim\frac{\sqrt{m}}{n^{\frac{k-2r+1}{2k}}}\le n^{\frac12-\frac{3}{2k}}\sqrt{m}. \tag{2.24}$$

Consequently,

$$
\begin{aligned}
|Y| &\overset{(2.20)}{\le} 2n|U| + 2|W| + 2|\mathsf{E}| + n \\
&\overset{(2.21)}{\lesssim} n\sum_{x\in[n]} \big|B_{\underline{d}}\big(x, h(x)-1\big)\big| + \sum_{u\in[n]} \big|B_{\underline{d}}\big(u, h(u)\big)\big| \deg_{\mathsf{G}}(u) + m \\
&\overset{(2.24)}{\lesssim} n\sum_{x\in[n]} \frac{\sqrt{m}}{n^{\frac{k-2(h(x)-1)+1}{2k}}} + n^{\frac12-\frac{3}{2k}}\sqrt{m}\sum_{u\in[n]}\deg_{\mathsf{G}}(u) + m \\
&\asymp n^{\frac12-\frac{3}{2k}}\sqrt{m}\sum_{x\in[n]} n^{\frac{h(x)}{k}} + n^{\frac12-\frac{3}{2k}} m^{\frac32} \\
&\overset{(2.7)}{\le} n^{\frac12-\frac{3}{2k}}\sqrt{m}\sum_{x\in[n]}\Big(1+\frac{1}{\sqrt{m}} n^{\frac{k+1}{2k}}\deg_{\mathsf{G}}(x)\Big) + n^{\frac12-\frac{3}{2k}} m^{\frac32} \\
&\asymp n^{\frac32-\frac{3}{2k}}\sqrt{m} + n^{1-\frac1k} m + n^{\frac12-\frac{3}{2k}} m^{\frac32} \\
&= o(n^2),
\end{aligned}
$$

where we used (twice) the fact that $\sum_{x\in[n]}\deg_{\mathsf{G}}(x) = 2m$ and the last step is valid because $m = o(n^{1+1/k})$.

3 Overview of the proof of Theorem 1.8 The Euclidean cone over a metric space $(\mathcal{M}, d_{\mathcal{M}})$, denoted $\mathsf{Cone}(\mathcal{M}, d_{\mathcal{M}})$ or $\mathsf{Cone}(\mathcal{M})$ when the metric is clear from the context, is defined [11] as the completion of $(0,\infty)\times\mathcal{M}$ under the following metric:

$$
(3.1)\qquad \forall (s,x),(t,y)\in(0,\infty)\times\mathcal{M},\qquad d_{\mathsf{Cone}(\mathcal{M},d_{\mathcal{M}})}\big((s,x),(t,y)\big) \overset{\mathrm{def}}{=} \sqrt{s^2+t^2-2st\cos\big(\min\{\pi, d_X(x,y)\}\big)}.
$$

See [1] and [15, Chapter I.5] for a treatment of this useful concept (including an explanation of the nomenclature); in particular, [15, Proposition I.5.9] proves that (3.1) indeed defines a metric.

The following general proposition is a key component of the proof of Theorem 1.8:

PROPOSITION 3.1. *For every metric transform $\varphi:[0,\infty)\to[0,\infty)$ there exist coefficients $\alpha_0 = \alpha_0^{\varphi}\ge 0$ and $\{\alpha_n = \alpha_n^{\varphi}\}_{n=1}^{\infty}, \{\beta_n = \beta_n^{\varphi}\}_{n=1}^{\infty} \subseteq (0,\infty)$ with the following property. Suppose that $(\mathcal{M}, d_{\mathcal{M}})$ is a metric space. There exist $\{z_k\}_{k=0}^{\infty}\subseteq\mathcal{M}$ and $\{r_k\}_{k=1}^{\infty}\subseteq(0,\infty)$ such that if we consider the Pythagorean product $(\mathcal{P},\rho)$, where*

$$
\mathcal{P} \overset{\mathrm{def}}{=} \left\{ \big(x_0,(s_1,x_1),(s_2,x_2),\ldots,\big)\in\mathcal{M}\times\big((0,\infty)\times\mathcal{M}\big)^{\mathbb{N}} :\ \sum_{k=1}^{\infty}\beta_k d_{\mathsf{Cone}\left(\mathcal{M},\pi\sqrt{\frac{\alpha_k}{\beta_k}}d_{\mathcal{M}}\right)}\big((s_k,x_k),(r_k,z_k)\big)^2<\infty \right\},
$$

and ρ is the metric on $\mathcal{P}\subseteq\mathcal{M}\times\big((0,\infty)\times\mathcal{M}\big)^{\mathbb{N}}$ that is given by setting

$$
\rho(\chi,\upsilon) \overset{\mathrm{def}}{=} \sqrt{\alpha_0 d_{\mathcal{M}}(x_0,y_0)^2 + \frac{1}{\pi^2}\sum_{k=1}^{\infty}\beta_k d_{\mathsf{Cone}\left(\mathcal{M},\pi\sqrt{\frac{\alpha_k}{\beta_k}}d_{\mathcal{M}}\right)}\big((s_k,x_k),(t_k,y_k)\big)^2},
$$

for each $\chi = \big(x_0,(s_1,x_1),(s_2,x_2),\ldots,\big), \upsilon = \big(y_0,(t_1,y_1),(t_2,y_2),\ldots,\big)\in\mathcal{P}$. Then $\mathsf{c}_{(\mathcal{P},\rho)}(\mathcal{M},\varphi\circ d_{\mathcal{M}})\lesssim 1$.

The full details of the proof of Theorem 3.1 appear in [27]; we will next sketch its steps. The following lemma shows that for any metric space $(\mathcal{M}, d_{\mathcal{M}})$, the truncated metric space $(\mathcal{M}, \min\{d_{\mathcal{M}},\pi\})$ embeds into the Euclidean cone over $(\mathcal{M}, d_{\mathcal{M}})$ with bi-Lipschitz distortion $O(1)$. The corresponding embedding is simply $(x\in\mathcal{M})\mapsto(1,x)\in(0,\infty)\times\mathcal{M}$, for which the stated distortion bound is proved in a straightforward way using elementary calculus; the details appear in [27].

LEMMA 3.2. *Every metric space $(\mathcal{M}, d_{\mathcal{M}})$ satisfies $\mathsf{c}_{\mathsf{Cone}(\mathcal{M},d_{\mathcal{M}})}\big(\mathcal{M},\min\{d_{\mathcal{M}},\pi\}\big)\le\frac{\pi}{2}$.*

The connection between Theorem 3.2 and Theorem 3.1 is obtained using the following representation by Brudnyĭ and Krugljak [17] of any metric transform $\omega:[0,\infty)\to[0,\infty)$ as an affine combination of truncations.

By [17, Proposition 3.2.6], there are $\alpha_0=\alpha_0^\omega\ge 0$ and $\{\alpha_n=\alpha_n^\omega\}_{n=1}^\infty,\{\beta_n=\beta_n^\omega\}_{n=1}^\infty\subseteq(0,\infty)$ such that

$$\forall t\ge 0,\qquad \omega(t)\asymp \alpha_0 t+\sum_{k=1}^\infty \min\{\alpha_k t,\beta_k\}.$$

Furthermore, $\lim_{t\to\infty}\omega(t)/t>0$ if and only if $\alpha_0>0$.

It is elementary to verify (as is done in, e.g., [49, Remark 5.4]) that if $\varphi:[0,\infty)\to[0,\infty)$ is a metric transform, then the mapping $\omega=\omega^\varphi:[0,\infty)\to[0,\infty)$, given by setting

$$\forall t\ge 0,\qquad \omega(t)\stackrel{\mathrm{def}}{=}\varphi\big(\sqrt{t}\big)^2$$

is also a metric transform. Using the aforementioned general representation of this specific ω, we see that

$$\forall t\ge 0,\qquad \varphi(t)\asymp\sqrt{\alpha_0 t^2+\sum_{k=1}^\infty \min\{\alpha_k t^2,\beta_k\}}. \tag{3.2}$$

Theorem 3.1 now follows by combining (3.2) with Theorem 3.2 through a sequence of elementary estimates that are performed in [27].

With Theorem 3.1 at hand, we can now outline how Theorem 1.8 is proved; the complete details of the subsequent deduction appear in [27].

A metric space $(\mathcal{N},d_\mathcal{N})$ is said to be a CAT(1) space if it satisfies the following conditions. Firstly, we require that for every $x,y\in\mathcal{N}$ with $d_\mathcal{N}(x,y)<\pi$ there exists a constant speed geodesic in $\mathcal{N}$ joining x to y. Next, suppose that $x,y,z\in\mathcal{N}$ satisfies $d_\mathcal{N}(x,y)+d_\mathcal{N}(y,z)+d_\mathcal{N}(z,x)<2\pi$, and that $\gamma_{x,y},\gamma_{y,z},\gamma_{z,x}:[0,1]\to\mathcal{N}$ are geodesics that join x to y, y to z, and z to x, respectively. Let $\mathbb{S}^2$ be the unit Euclidean sphere in $\mathbb{R}^3$, and let $d_{\mathbb{S}^2}$ denote the geodesic metric on $\mathbb{S}^2$ (thus the diameter of $\mathbb{S}^2$ equals π under this metric). As explained in [15], there exist $a,b,c\in\mathbb{S}^2$ such that $d_{\mathbb{S}^2}(a,b)=d_\mathcal{N}(x,y)$, $d_{\mathbb{S}^2}(b,c)=d_\mathcal{N}(y,z)$, and $d_{\mathbb{S}^2}(c,a)=d_\mathcal{N}(z,x)$. Let $\phi_{a,b},\phi_{b,c},\phi_{c,a}:[0,1]\to\mathbb{S}^2$ be $d_{\mathbb{S}^2}$-geodesics that join a to b, b to c, and c to a, respectively. The second requirement in the definition of a CAT(1) space is that $d_\mathcal{N}(\phi_{x,y}(s),\phi_{y,z}(t))\le d_{\mathbb{S}^2}(\phi_{a,b}(s),\phi_{b,c}(t))$ for every $s,t\in[0,1]$. See [19, 65, 15, 70, 18, 71, 35] for more on this fundamental notion.

A straightforward consequence of the relation between distances in $\mathbb{R}^2$ and $\mathbb{S}^2$ shows that every CAT(0) space is also a CAT(1) space; see [15, Theorem II.1.12]. An important theorem of Berestovskiĭ [11] (see also [1] and [15, Theorem II.3.14]) states that a metric space $(\mathcal{N},d_\mathcal{N})$ is a CAT(1) space if and only if $\mathsf{Cone}(\mathcal{N},d_\mathcal{N})$ is a CAT(0) space. By combining these two facts (and the obvious fact that if one multiplies the metric of a CAT(0) space by a positive constant, then the resulting metric space is also a CAT(0) space), if $(\mathcal{M},d_\mathcal{M})$ is a CAT(0) metric space, then $\mathsf{Cone}(\mathcal{M},sd_\mathcal{M})$ is also a CAT(0) space for any $s>0$.

Returning to the setting of Theorem 1.8, we are given a CAT(0) metric space $(\mathcal{M},d_\mathcal{M})$ and a metric transform $\varphi:[0,\infty)\to[0,\infty)$. We will next apply Theorem 3.1, and proceed using the notation in its statement. As explained above, $\mathsf{Cone}(\mathcal{M},\pi\sqrt{\alpha_k/\beta_k}d_\mathcal{M})$ is a CAT(0) space for each $k\in\mathbb{N}$. Consequently, the Pythagorean product $(\mathcal{P},\rho)$ of Theorem 3.1 is also a CAT(0) space, since the CAT(0) condition (1.7) is a quadratic inequality that evidently passes to Pythagorean products (by applying it coordinate-wise). Theorem 3.1 asserts that the metric space $(\mathcal{M},\varphi\circ d_\mathcal{M})$ admits an embedding into the CAT(0) space $(\mathcal{P},\rho)$ whose bi-Lipschitz distortion is $O(1)$, so the conclusion of Theorem 1.8 holds.

Added in proof The concurrent work [3] made substantial (exciting) progress towards Question 1.9 by proving that, with high probability, a random regular graph is a universal $O(1)$-approximator with respect to an independently chosen random regular graph. This resolves the question of Kleinberg that Theorem 1.9 mentions and provides a randomized algorithm that performs as described in Theorem 1.6. Derandomizing that construction remains a major challenge; see [3] for more details.

References

[1] A. D. Aleksandrov, V. N. Berestovskiĭ, and I. G. Nikolaev, *Generalized Riemannian spaces*, Uspekhi Mat. Nauk, 41 (1986), pp. 3–44, 240.

[2] I. Althöfer, G. Das, D. Dobkin, D. Joseph, and J. Soares, *On sparse spanners of weighted graphs*, Discrete Comput. Geom., 9 (1993), pp. 81–100, https://doi.org/10.1007/BF02189308.

[3] D. J. Altschuler, P. Dodos, K. Tikhomirov, and K. Tyros, *Discrete poincaré inequalities and universal approximators for random graphs*, 2025, https://arxiv.org/abs/2506.17433.

[4] A. Andoni, A. Naor, and O. Neiman, *Snowflake universality of Wasserstein spaces*, Ann. Sci. Éc. Norm. Supér. (4), 51 (2018), pp. 657–700, https://doi.org/10.24033/asens.2363.

[5] A. Andoni, A. Naor, A. Nikolov, I. Razenshteyn, and E. Waingarten, *Data-dependent hashing via nonlinear spectral gaps*, in STOC'18—Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, ACM, New York, 2018, pp. 787–800, https://doi.org/10.1145/3188745.3188846, https://doi.org/10.1145/3188745.3188846.

[6] K. Ball, *Markov chains, Riesz transforms and Lipschitz maps*, Geom. Funct. Anal., 2 (1992), pp. 137–172, https://doi.org/10.1007/BF01896971.

[7] K. Ball, *The Ribe programme*, Astérisque, (2013), pp. Exp. No. 1047, viii, 147–159.

[8] I. Bárány and Z. Füredi, *Computing the volume is difficult*, Discrete Comput. Geom., 2 (1987), pp. 319–326, https://doi.org/10.1007/BF02187886.

[9] K. Barhum, O. Goldreich, and A. Shraibman, *On approximating the average distance between points*, in 10th International Workshop, APPROX 2007, and 11th International Workshop, RANDOM 2007, vol. 4627 of LNCS, Springer, 2007, pp. 296–310, https://doi.org/10.1007/978-3-540-74208-1_22.

[10] M. Bačák, *Convex analysis and optimization in Hadamard spaces*, vol. 22 of De Gruyter Series in Nonlinear Analysis and Applications, De Gruyter, Berlin, 2014, https://doi.org/10.1515/9783110361629, https://doi.org/10.1515/9783110361629.

[11] V. N. Berestovskiĭ, *Borsuk's problem on metrization of a polyhedron*, Dokl. Akad. Nauk SSSR, 268 (1983), pp. 273–277.

[12] L. M. Blumenthal, *Concerning spherical spaces*, American Journal of Mathematics, 57 (1935), pp. 51–61.

[13] L. M. Blumenthal, *Theory and applications of distance geometry*, Oxford, at the Clarendon Press,, 1953.

[14] J. Bourgain, *The metrical interpretation of superreflexivity in Banach spaces*, Israel J. Math., 56 (1986), pp. 222–230, https://doi.org/10.1007/BF02766125.

[15] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, vol. 319, Springer-Verlag, Berlin, 1999, https://doi.org/10.1007/978-3-662-12494-9, https://doi.org/10.1007/978-3-662-12494-9.

[16] A. Brudnyi and Y. Brudnyi, *Methods of geometric analysis in extension and trace problems. Volume 2*, vol. 103 of Monographs in Mathematics, Birkhäuser/Springer Basel AG, Basel, 2012.

[17] Y. A. Brudnyĭ and N. Y. Krugljak, *Interpolation functors and interpolation spaces. Vol. I*, vol. 47 of North-Holland Mathematical Library, North-Holland Publishing Co., Amsterdam, 1991. Translated from the Russian by Natalie Wadhwa, With a preface by Jaak Peetre.

[18] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, vol. 33 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2001, https://doi.org/10.1090/gsm/033, https://doi.org/10.1090/gsm/033.

[19] Y. Burago, M. Gromov, and G. Perelman, *A. D. Aleksandrov spaces with curvatures bounded below*, Uspekhi Mat. Nauk, 47 (1992), pp. 3–51, 222, https://doi.org/10.1070/RM1992v047n02ABEH000877.

[20] T. de Laat and M. de la Salle, *Banach space actions and L^2-spectral gap*, Anal. PDE, 14 (2021), pp. 45–76, https://doi.org/10.2140/apde.2021.14.45.

[21] M. M. Deza and M. Laurent, *Geometry of cuts and metrics*, vol. 15 of Algorithms and Combinatorics, Springer, Heidelberg, 2010, https://doi.org/10.1007/978-3-642-04295-9, https://doi.org/10.1007/978-3-642-04295-9. First softcover printing of the 1997 original [MR1460488].

[22] J. Ding, J. R. Lee, and Y. Peres, *Markov type and threshold embeddings*, Geom. Funct. Anal., 23 (2013), pp. 1207–1229, https://doi.org/10.1007/s00039-013-0234-7.

[23] M. Dyer, A. Frieze, and R. Kannan, *A random polynomial-time algorithm for approximating the volume of convex bodies*, J. Assoc. Comput. Mach., 38 (1991), pp. 1–17, https://doi.org/10.1145/102782.102783.

[24] P. Erdős, *Extremal problems in graph theory*, in Theory of Graphs and its Applications (Proc. Sympos. Smolenice, 1963), Publ. House Czech. Acad. Sci., Prague, 1964, pp. 29–36.

[25] A. Eskenazis, *Average distortion embeddings, nonlinear spectral gaps, and a metric john theorem [after Assaf Naor]*, Astérisque, (2022), pp. 295–333.

[26] A. Eskenazis, M. Mendel, and A. Naor, *Nonpositive curvature is not coarsely universal*, Invent. Math., 217 (2019), pp. 833–886, https://doi.org/10.1007/s00222-019-00878-1.

[27] A. Eskenazis, M. Mendel, and A. Naor, *Hadamard subspaces are closed under metric transforms*, 2025. Preprint available at [28, Part 2].

[28] A. Eskenazis, M. Mendel, and A. Naor, *An optimal algorithm for average distance in typical regular graphs*, 2025, https://arxiv.org/abs/2510.18722v1.

[29] E. D. Gluskin, *The diameter of the Minkowski compactum is roughly equal to n*, Funktsional. Anal. i Prilozhen., 15 (1981), pp. 72–73.

[30] G. Godefroy, *De Grothendieck à Naor: une promenade dans l'analyse métrique des espaces de Banach*, Gaz. Math., (2017), pp. 13–24.

[31] O. Goldreich and D. Ron, *Approximating average parameters of graphs*, in Approximation, randomization and combinatorial optimization, vol. 4110 of Lecture Notes in Comput. Sci., Springer, Berlin, 2006, pp. 363–374, https://doi.org/10.1007/11830924_34, https://doi.org/10.1007/11830924_34.

[32] O. Goldreich and D. Ron, *Approximating average parameters of graphs*, Random Structures Algorithms, 32 (2008), pp. 473–493, https://doi.org/10.1002/rsa.20203.

[33] M. Gromov, *Hyperbolic groups*, in Essays in group theory, vol. 8 of Math. Sci. Res. Inst. Publ., Springer, New York, 1987, pp. 75–263, https://doi.org/10.1007/978-1-4613-9586-7_3, https://doi.org/10.1007/978-1-4613-9586-7_3.

[34] M. Gromov, *Random walk in random groups*, Geom. Funct. Anal., 13 (2003), pp. 73–146, https://doi.org/10.1007/s000390300002.

[35] M. Gromov, *Metric structures for Riemannian and non-Riemannian spaces*, Modern Birkhäuser Classics, Birkhäuser Boston, Inc., Boston, MA, english ed., 2007. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates.

[36] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization*, vol. 2 of Algorithms and Combinatorics, Springer-Verlag, Berlin, second ed., 1993, https://doi.org/10.1007/978-3-642-78240-4, https://doi.org/10.1007/978-3-642-78240-4.

[37] P. Indyk, *Sublinear time algorithms for metric space problems*, in Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999), ACM, New York, 1999, pp. 428–432, https://doi.org/10.1145/301250.301366, https://doi.org/10.1145/301250.301366.

[38] H. Izeki and S. Nayatani, *Combinatorial harmonic maps and discrete-group actions on Hadamard spaces*, Geom. Dedicata, 114 (2005), pp. 147–188, https://doi.org/10.1007/s10711-004-1843-y.

[39] J. Jost, *Nonpositive curvature: geometric and analytic aspects*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1997, https://doi.org/10.1007/978-3-0348-8918-6, https://doi.org/10.1007/978-3-0348-8918-6.

[40] N. J. Kalton, *The nonlinear geometry of Banach spaces*, Rev. Mat. Complut., 21 (2008), pp. 7–60, https://doi.org/10.5209/rev_REMA.2008.v21.n1.16426.

[41] T. Kondo, *Cat(0) spaces and expanders*, Math. Z., 271 (2012), pp. 343–355, https://doi.org/10.1007/s00209-011-0866-y.

[42] S. Kwapień, *Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients*, Studia Math., 44 (1972), pp. 583–595, https://doi.org/10.4064/sm-44-6-583-595.

[43] V. Lafforgue, *Un renforcement de la propriété (T)*, Duke Math. J., 143 (2008), pp. 559–602, https://doi.org/10.1215/00127094-2008-029.

[44] J. R. Lee, A. Naor, and Y. Peres, *Trees and Markov convexity*, Geom. Funct. Anal., 18 (2009), pp. 1609–1659, https://doi.org/10.1007/s00039-008-0689-0.

[45] N. Linial, E. London, and Y. Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Combinatorica, 15 (1995), pp. 215–245, https://doi.org/10.1007/BF01200757.

[46] J. Lurie, *Some notes that i once prepared on the theory of hadamard spaces*, https://www.math.ias.edu/~lurie/papers/hadamard.pdf.

[47] J. Matoušek, *Lectures on discrete geometry*, vol. 212 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2002, https://doi.org/10.1007/978-1-4613-0039-7, http://dx.doi.org/10.1007/978-1-4613-0039-7.

[48] J. Matoušek, *On embedding expanders into l_p spaces*, Israel J. Math., 102 (1997), pp. 189–197, https://doi.org/10.1007/BF02773799.

[49] M. Mendel and A. Naor, *Euclidean quotients of finite metric spaces*, Adv. Math., 189 (2004), pp. 451–494, https://doi.org/10.1016/j.aim.2003.12.001.

[50] M. Mendel and A. Naor, *Metric cotype*, Ann. of Math. (2), 168 (2008), pp. 247–298, https://doi.org/10.4007/annals.2008.168.247.

[51] M. Mendel and A. Naor, *Spectral calculus and Lipschitz extension for barycentric metric spaces*, Anal. Geom. Metr. Spaces, 1 (2013), pp. 163–199, https://doi.org/10.2478/agms-2013-0003.

[52] M. Mendel and A. Naor, *Expanders with respect to Hadamard spaces and random graphs [extended abstract]*, in ITCS'14—Proceedings of the 2014 Conference on Innovations in Theoretical Computer Science, ACM, New York, 2014, pp. 353–358.

[53] M. Mendel and A. Naor, *Nonlinear spectral calculus and super-expanders*, Publ. Math. Inst. Hautes Études Sci., 119 (2014), pp. 1–95, https://doi.org/10.1007/s10240-013-0053-2.

[54] M. Mendel and A. Naor, *Expanders with respect to Hadamard spaces and random graphs*, Duke Math. J., 164 (2015), pp. 1471–1548, https://doi.org/10.1215/00127094-3119525.

[55] M. Mimura, *Sphere equivalence, Banach expanders, and extrapolation*, Int. Math. Res. Not. IMRN, (2015), pp. 4372–4391, https://doi.org/10.1093/imrn/rnu075.

[56] A. Naor, *An introduction to the ribe program*, Japanese Journal of Mathematics, 7 (2012), pp. 167–233, https://doi.org/10.1007/s11537-012-1222-7.

[57] A. Naor, *Comparison of metric spectral gaps*, Anal. Geom. Metr. Spaces, 2 (2014), pp. 1–52, https://doi.org/10.2478/agms-2014-0001.

[58] A. Naor, *Metric dimension reduction: a snapshot of the Ribe program*, in Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. I. Plenary lectures, World Sci. Publ., Hackensack, NJ, 2018, pp. 759–837.

[59] A. Naor, *An average John theorem*, Geom. Topol., 25 (2021), pp. 1631–1717, https://doi.org/10.2140/gt.2021.25.1631.

[60] A. Naor, *Mixing convexity*. Preprint, 2025.

[61] A. Naor and L. Silberman, *Poincaré inequalities, embeddings, and wild groups*, Compos. Math., 147 (2011), pp. 1546–1572, https://doi.org/10.1112/S0010437X11005343.

[62] P. W. Nowak and G. Yu, *Large scale geometry*, EMS Textbooks in Mathematics, European Mathematical Society (EMS), Zürich, 2012, https://doi.org/10.4171/112, https://doi.org/10.4171/112.

[63] M. I. Ostrovskii, *Metric embeddings*, vol. 49 of De Gruyter Studies in Mathematics, De Gruyter, Berlin, 2013, https://doi.org/10.1515/9783110264012, https://doi.org/10.1515/9783110264012. Bilipschitz and coarse embeddings into Banach spaces.

[64] N. Ozawa, *A note on non-amenability of $B(l_p)$ for $p = 1, 2$*, Internat. J. Math., 15 (2004), pp. 557–565, https://doi.org/10.1142/S0129167X04002430.

[65] G. Perelman, *Spaces with curvature bounded below*, in Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994), Birkhäuser, Basel, 1995, pp. 517–525.

[66] G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math., 20 (1975), pp. 326–350, https://doi.org/10.1007/BF02760337.

[67] G. Pisier, *Complex interpolation between Hilbert, Banach and operator spaces*, Mem. Amer. Math. Soc., 208 (2010), pp. vi+78, https://doi.org/10.1090/S0065-9266-10-00601-0.

[68] O. Reingold, S. Vadhan, and A. Wigderson, *Entropy waves, the zig-zag graph product, and new constant-degree expanders*, Ann. of Math. (2), 155 (2002), pp. 157–187, https://doi.org/10.2307/3062153.

[69] J. Roe, *Lectures on coarse geometry*, vol. 31 of University Lecture Series, American Mathematical Society, Providence, RI, 2003, https://doi.org/10.1090/ulect/031, https://doi.org/10.1090/ulect/031.

[70] K. T. Sturm, *Metric spaces of lower bounded curvature*, Exposition. Math., 17 (1999), pp. 35–47.

[71] K.-T. Sturm, *Probability measures on metric spaces of nonpositive curvature*, in Heat kernels and analysis on manifolds, graphs, and metric spaces (Paris, 2002), vol. 338 of Contemp. Math., Amer. Math. Soc., Providence, RI, 2003, pp. 357–390, https://doi.org/10.1090/conm/338/06080, https://doi.org/10.1090/conm/338/06080.

[72] S. J. Szarek and N. Tomczak-Jaegermann, *Saturating constructions for normed spaces*, Geom. Funct. Anal., 14 (2004), pp. 1352–1375, https://doi.org/10.1007/s00039-004-0495-2.

[73] M. Thorup and U. Zwick, *Approximate distance oracles*, J. ACM, 52 (2005), pp. 1–24, https://doi.org/10.1145/1044731.1044732.

[74] M.-T. Wang, *A fixed point theorem of discrete group actions on Riemannian manifolds*, J. Differential Geom., 50 (1998), pp. 249–267.

[75] J. H. Wells and L. R. Williams, *Embeddings and extensions in analysis*, vol. Band 84 of Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], Springer-Verlag, New York-Heidelberg, 1975.

[76] W. A. Wilson, *A Relation Between Metric and Euclidean Spaces*, Amer. J. Math., 54 (1932), pp. 505–517, https://doi.org/10.2307/2370894.