ASSOUAD'S THEOREM WITH DIMENSION INDEPENDENT OF THE SNOWFLAKING

ASSAF NAOR AND OFER NEIMAN

Abstract. It is shown that for every $K > 0$ and $\varepsilon \in (0, 1/2)$ there exist $N = N(K) \in \mathbb{N}$ and $D = D(K, \varepsilon) \in (1, \infty)$ with the following properties. For every metric space (X, d) with doubling constant at most K, the metric space $(X, d^{1-\varepsilon})$ admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion at most D. The classical Assouad embedding theorem makes the same assertion, but with $N \to \infty$ as $\varepsilon \to 0$.

1. Introduction

In this paper all metric spaces are assumed to be separable and contain at least two points. Balls in metric spaces are always closed balls, i.e., for a metric space (X, d), $x \in X$ and $r \geqslant 0$, we denote $B(x, r) = \{y \in X : \ d(x, y) \leqslant r\}$. A metric space (X, d) has doubling constant $K \in (1, \infty)$ if every ball in X can be covered by at most K balls of half its radius, i.e., for every $x \in X$ and $r > 0$ there exist $A \subseteq X$ with $|A| \leqslant K$ such that $B(x, r) \subseteq \bigcup_{y \in A} B(y, r/2)$. Note that since X contains at least two points, necessarily $K \geqslant 2$. (X, d) is said to be a doubling metric space if it has doubling constant K for some $K \in (1, \infty)$.

A metric space (X, d) embeds into a normed space $(Y, \|\cdot\|)$ with distortion $D \in [1, \infty]$ if there exists $f : X \to Y$ such that for all $x, y \in X$ we have $d(x, y) \leqslant \|f(x) - f(y)\| \leqslant Dd(x, y)$. When X embeds into Y with finite distortion we say that X admits a bi-Lipschitz embedding into Y. The infimum over those $D \geqslant 1$ for which X embeds into Y is denoted $c_Y(X)$. When $Y = \ell_2$ is infinite dimensional Hilbert space, we write $c_Y(X) = c_2(X)$; this parameter is known in the literature as the Euclidean distortion of X. In what follows, when we refer to the space $\mathbb{R}^N$ we always assume that it is equipped with the standard Euclidean metric. A standard argument (see, e.g., [4, Lem. 4.9]) shows that if Y is either ℓ_2 or $\mathbb{R}^N$, we have

$$c_Y(X) = \sup\{c_Y(Z) : \ Z \subseteq X \ \wedge \ |Z| < \infty\}. \tag{1}$$

If (X, d) is a metric space and $\alpha \in (0, 1]$ then (X, d^α) is also a metric space, known as the α-snowflake of X.

A major open problem of embedding theory is the bi-Lipschitz embeddability problem in $\mathbb{R}^N$. This problem asks for an intrinsic characterization of those separable metric spaces (X, d) that admit a bi-Lipschitz embedding into $\mathbb{R}^N$ for some $N \in \mathbb{N}$. For a discussion of this important question, see for example the works of Semmes [28], Lang-Plaut [20] and Heinonen [16]. An obvious restriction on a metric space (X, d) that admits a bi-Lipschitz embedding into $\mathbb{R}^N$ is that it must be doubling. In this context, Assouad discovered in [4] the following fundamental embedding theorem (see also Heinonen's book [15] for a nice exposition of Assouad's theorem).

A. N. was supported in part by NSF grant CCF-0832795, BSF grant 2006009, and the Packard Foundation. O. N. was supported in part by NSF grant CCF-0832795.

Theorem 1.1 (Assouad's embedding theorem). *For every $\varepsilon \in (0,1)$ and $K > 0$, there exist $N = N(K,\varepsilon) \in \mathbb{N}$ and $D = D(K,\varepsilon) \in (1,\infty)$ such that for every separable metric space (X,d) with doubling constant K, the metric space $(X,d^{1-\varepsilon})$ admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion at most D.*

Assouad's theorem falls short of solving the bi-Lipschitz embeddability problem in $\mathbb{R}^N$, since it only achieves an embedding of the snowflaked metric space $(X,d^{1-\varepsilon})$. Nevertheless, as $\varepsilon \to 0$ this metric space becomes closer and closer to the original metric space (X,d). It is therefore of interest to investigate the behavior of $N(K,\varepsilon)$ and $D(K,\varepsilon)$ as $\varepsilon \to 0$. It turns out that necessarily $\lim_{\varepsilon\to 0} D(K,\varepsilon) = \infty$, due to the existence of doubling metric spaces that do not admit a bi-Lipschitz embedding into $\mathbb{R}^N$. The first known such example is the Heisenberg group, equipped with the Carnot-Carathéodory metric: Semmes observed in [27] that its bi-Lipschitz nonembeddability into $\mathbb{R}^N$ is a consequence of Pansu's differentiability theorem [25]. Additional examples of non-Euclidean doubling spaces were found by Laakso [19] and Bourdon-Pajot [7]; see the work of Cheeger [8] for a unified treatment of these results.

It seems to be inherent to Assouad's embedding method that also $\lim_{\varepsilon\to 0} N(K,\varepsilon) = \infty$. Note that if $\varepsilon \in (0,1/2)$ then the metric space $(X,d^{1-\varepsilon})$ has doubling constant K^2, so there is no obvious obstruction to $(X,d^{1-\varepsilon})$ admitting a bi-Lipschitz embedding into $\mathbb{R}^N$ for some N that is independent of $\varepsilon \in (0,1/2)$. The issue that in Assouad's theorem N depends on ε and is very large as $\varepsilon \to 0$ was noted by many authors; this is mentioned, for example, in the works of David-Toro [10] and Semmes [28] (where much more refined bounds on N are obtained under additional assumptions). Assouad himself noticed this issue in [4], where he showed that N can be taken to be independent of $\varepsilon \in (0,1/2)$ when $X = \mathbb{R}$ (more generally, Assouad deals in [4] with $X = [0,1]^k$). The case of the "helix snowflakes" $(\mathbb{R}, |x-y|^{1-\varepsilon})$ was studied by Kahane [17] and Talagrand [29], who investigated the interplay between the dimension N and the distortion D (Kahane studied only the case $\varepsilon = \frac{1}{2}$, and obtained sharp results. Talagrand's work applies to all $\varepsilon \in (0,1)$, but is not sharp).

Here we show that in Assouad's theorem one can take N to depend only on the doubling constant K, but not on $\varepsilon \in (0,1/2)$.

Theorem 1.2. *For every $K > 0$ there exists $N = N(K) \in \mathbb{N}$, and for every $\varepsilon \in (0,1/2)$ and $K > 0$ there exists $D = D(K,\varepsilon) \in (1,\infty)$, such that for every separable metric space (X,d) with doubling constant K, the metric space $(X,d^{1-\varepsilon})$ admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion at most D.*

Our argument yields the bounds $N(K) \lesssim \log K$ and $D(K,\varepsilon) \lesssim \left(\frac{\log K}{\varepsilon}\right)^2$. More generally, for every $\delta \in (0,1]$ our argument yields the bounds

$$N(K) \lesssim \frac{\log K}{\delta} \quad \text{and} \quad D(K,\varepsilon) \lesssim \left(\frac{\log K}{\varepsilon}\right)^{1+\delta}. \tag{2}$$

Here and in what follows, the symbols $\lesssim, \gtrsim$ indicate the corresponding inequalities up to an absolute multiplicative factor.

In the rest of this introduction we will describe some additional results and question related to the bi-Lipschitz embeddability problem in $\mathbb{R}^N$.

1.1. **The Lang-Plaut problem and snowflakes of the Heisenberg group.** Despite major efforts by many mathematicians, the bi-Lipschitz embeddability problem in $\mathbb{R}^N$ remains wide open. A variety of sufficient intrinsic conditions on a metric space (X,d) are known which ensure that it admits a bi-Lipschitz embedding in some Euclidean space $\mathbb{R}^N$, but these conditions are far from necessary.

A necessary condition for a metric space (X,d) to admit a bi-Lipschitz embedding into some $\mathbb{R}^N$ (in addition to being doubling) is that it admits a bi-Lipschitz embedding into ℓ_2, i.e., its Euclidean distortion satisfies $c_2(X)<\infty$. All the known examples of doubling metric spaces that do not admit a bi-Lipschitz embedding into any $\mathbb{R}^N$ actually do not admit a bi-Lipschitz embedding into infinite dimensional Hilbert space as well. This led Lang and Plaut [20, Question 2.4] to ask the following question.

Question 1 (Lang-Plaut problem)**.** *Is it necessary and sufficient for a metric space (X,d) to admit a bi-Lipschitz embedding into some $\mathbb{R}^N$ that it is doubling and it admits a bi-Lipschitz embedding into Hilbert space? Equivalently, does every doubling subset of Hilbert space admit a bi-Lipschitz embedding into some $\mathbb{R}^N$?*

By a simple argument (presented in Section 4), the Lang-Plaut problem can be restated quantitatively as follows. Is it true that for every $K>0$ there is $N=N(K)\in\mathbb{N}$ and $D=D(K)\in(1,\infty)$ such that if $X\subseteq\ell_2$ has doubling constant K then $c_{\mathbb{R}^N}(X)\leqslant D$?

One might argue whether or not a positive answer to the Lang-Plaut problem would resolve the bi-Lipschitz embedding problem into $\mathbb{R}^N$, since it is not obvious that the condition that X admits a bi-Lipschitz embedding into Hilbert space can be restated in terms of the intrinsic geometry of X. But, it is possible to characterize bi-Lipschitz embeddability into ℓ_2 in terms of a family of distance inequalities, i.e., intrinsically, without using the word "embedding". Indeed, as shown by Linial-London-Rabinovich [23] (extending the corresponding classical result of Schoenberg [26] in the isometric category), $c_2(X)\leqslant D$ if and only if for all $n\in\mathbb{N}$, $x_1,\ldots,x_n\in X$ and every $n\times n$ symmetric positive semidefinite matrix $Q=(q_{ij})$, all of whose rows sum to 0, the following inequality holds true:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\max\{q_{ij},0\}d(x_i,x_j)^2\leqslant D^2\sum_{i=1}^{n}\sum_{j=1}^{n}\max\{-q_{ij},0\}d(x_i,x_j)^2. \tag{3}$$

Hence, a positive answer to the Lang-Plaut question would yield a characterization of bi-Lipschitz embeddability into some $\mathbb{R}^N$ in terms of the doubling condition, and the family of distance inequalities (3). We believe that this would yield a satisfactory answer to the bi-Lipschitz embeddability problem in $\mathbb{R}^N$, though there does not seem to be evidence supporting a positive answer to the the Lang-Plaut question.

A potential source of doubling subsets of Hilbert space that might yield a counter-example to the Lang-Plaut problem is Assouad's theorem itself. When allowing embeddings into infinite dimensional Hilbert space rather than into $\mathbb{R}^N$, the asymptotics in terms of ε of $D(K,\varepsilon)$ in Assouad's theorem are known [21] (see also [24]). Specifically, if (X,d) has doubling constant K then $c_2(X)\leqslant C(K)/\sqrt{\varepsilon}$ for some $C(K)\in(0,\infty)$. This dependence on ε is sharp up to the value of $C(K)$, as shown in [21, Remark 5.4].

If (X,d) has doubling constant K then the space $(X,d^{1-\varepsilon})$ has doubling constant bounded uniformly in $\varepsilon\in(0,1/2)$, but, in its $C(K)/\sqrt{\varepsilon}$-distortion embedding into ℓ_2 it might have an image that is not a doubling subset of ℓ_2, with doubling constant independent of ε, due to the large distortion. We therefore ask the following question:

Question 2. *Is it true that for every $K \in (1,\infty)$ there exist $a(K), b(K) \in (0,\infty)$ with the following property. If (X,d) has doubling constant K and $\varepsilon \in (0,1/2)$ then there exists $f : X \to \ell_2$ such that $a(K)\sqrt{\varepsilon}d(x,y)^{1-\varepsilon} \leqslant \|f(x) - f(y)\| \leqslant d(x,y)^{1-\varepsilon}$ for all $x,y \in X$, and $f(X) \subseteq \ell_2$ has doubling constant $b(K)$.*

Observe that due to Theorem 1.2, with the explicit bounds stated in (2), if we replaced in Question 2 the term $\sqrt{\varepsilon}$ by $\varepsilon^{1+\delta}$ for any $\delta \in (0,1]$, then the answer would be positive, and even the image of the embedding would be finite dimensional with dimension depending only on K and δ.

In spite of the fact that we don't know the answer to Question 2, we do know that the answer is positive for the Heisenberg group. For $n \in \mathbb{N}$, the n'th Heisenberg group $\mathbb{H}_n$ is $\mathbb{C}^n \times \mathbb{R}$, equipped with the following group product:

$$(w,s)\cdot(z,t) = \left(w+z, s+t+2\sum_{j=1}^{n}\Im\left(w_j\overline{z_j}\right)\right)$$
$$\forall w = (w_1,\ldots,w_n), z = (z_1,\ldots,z_n) \in \mathbb{C}^n,\ \forall s,t \in \mathbb{R}.$$

Thus $(0,0)$ is the identity of $\mathbb{H}_n$ and for $(z,t) \in \mathbb{H}_n$ we have $(z,t)^{-1} = (-z,-t)$.

The Koranyi norm on $\mathbb{H}_n$ is defined for $(z,t) \in \mathbb{H}_n$ by $N_0(z,t) = \sqrt[4]{|z|^4 + t^2}$, where $|z|^2 = \sum_{j=1}^n |z_j|^2$. For $g,h \in \mathbb{H}_n$ we have $N_0(gh^{-1}) \leqslant N_0(g) + N_0(h)$ (see [18, 9]). Thus $d_{N_0}(g,h) = N_0(h^{-1}g)$ is a left-invariant metric on $\mathbb{H}_n$. One can check that the Lebesgue measure is a Haar measure of $\mathbb{H}_n$, and that $(\mathbb{H}_n, d_{N_0})$ has doubling constant $e^{O(n)}$.

In Section 4 we observe that a result of [22] implies the following statement.

Theorem 1.3. *For every $\varepsilon \in (0,1/2)$ and for every $n \in \mathbb{N}$, there exists $f_\varepsilon : \mathbb{H}_n \to \ell_2$ satisfying $\sqrt{\varepsilon}d_{N_0}(x,y)^{1-\varepsilon} \leqslant \|f_\varepsilon(x) - f_\varepsilon(y)\| \leqslant d_{N_0}(x,y)^{1-\varepsilon}$ for all $x,y \in \mathbb{H}_n$, and such that $f_\varepsilon(\mathbb{H}_n)$ is a doubling subset of ℓ_2, with doubling constant $e^{O(n)}$.*

We also show in Section 4 that Theorem 1.3 is sharp, even without the requirement that the image of $\mathbb{H}_n$ is doubling with constant independent of ε:

$$c_2\left(\mathbb{H}_n, d_{N_0}^{1-\varepsilon}\right) \gtrsim \frac{1}{\sqrt{\varepsilon}} \quad \forall\, \varepsilon \in (0,1/2). \tag{4}$$

This raises the following question:

Question 3. *Is it true that for every fixed $N \in \mathbb{N}$ we have $\lim_{\varepsilon\to 0} c_{\mathbb{R}^N}\left(\mathbb{H}_1, d_{N_0}^{1-\varepsilon}\right)\sqrt{\varepsilon} = \infty$?*

A positive answer to Question 3 would imply a negative answer to the Lang-Plaut problem, since otherwise there would be $N \in \mathbb{N}$ and $D \in (1,\infty)$ satisfying $c_{\mathbb{R}^N}(f_\varepsilon(\mathbb{H}_1)) \leqslant D$ for all $\varepsilon \in (0,1/2)$, where f_ε is the Euclidean embedding of $\left(\mathbb{H}_1, d_{N_0}^{1-\varepsilon}\right)$ from Theorem 1.3. This would yield the bound $c_{\mathbb{R}^N}\left(\mathbb{H}_1, d_{N_0}^{1-\varepsilon}\right)\sqrt{\varepsilon} \leqslant D$.

1.2. **Previous work and an overview of the proof of Theorem 1.2.** The classical proof of Assouad's theorem [4, 15] yields the dimension bound $N(K,\varepsilon) \leqslant c(K)/\varepsilon^{O(1)}$. In [13] Gupta-Krauthgamer-Lee announced a similar bound on $N(K,\varepsilon)$ with a much better dependence of $c(K)$ on K, yet the same bound in terms of ε (the proof of this assertion of [13] hasn't appeared since the 2003 announcement, and in particular the dependence on ε was not stated there explicitly, but it seems to us that the proof technique suggested in [13] would lead to this bound). A similar bound follows from the work of Har-Peled and Mendel [14], who studied in addition embeddings into ℓ_∞^N, yielding a $1+\delta$ distortion result. The best previously known bound is due to Abraham-Bartal-Neiman [1], who proved that $N(K,\varepsilon) \leqslant c(K)\log(1/\varepsilon)$. In the context of the Lang-Plaut problem, Gottlieb-Krauthgamer [12], and Bartal-Recht-Schulman [6], proved that if $X \subseteq \ell_2$ has doubling constant K then for all $\delta \in (0,1)$ the $(1-\varepsilon)$-snowflake of X embeds with distortion $1+\delta$ into $\mathbb{R}^{c(K,\delta)/\varepsilon^{O(1)}}$; the main point in these works, however, is to obtain a $1+\delta$ distortion embedding, which is impossible in the context of general doubling metric spaces that are not necessarily isometric to a subset of ℓ_2.

Our proof of Theorem 1.2 builds heavily on the method of Abraham-Bartal-Neiman [1]. In essence, our proof should be viewed as an optimization of the argument of [1] which uses degrees of freedom that were available in the construction of [1] but were not previously exploited. This requires subtle changes in the proof of [1], and in particular we were surprised that such changes can lead to a complete removal of the dependence on ε of the dimension N in Assouad's theorem. Though somewhat delicate, these changes are of a technical nature, and the key conceptual ideas can be found in [1].

The proof of Theorem 1.2 is based on a construction of a distribution over random embeddings, arising from a certain family of random multi-scale partitions of the metric space (X,d). At every possible distance scale we provide a mapping to $\mathbb{R}$ which is essentially the truncated distance to the "boundary" of the random partition. We then combine all the possible scales into a single embedding into $\mathbb{R}$, using an idea of Assouad [4] which multiplies every scale by an appropriate factor that enables us to control the total expansion over all scales. The lower bound on the distance of the image of every pair of points in X will come from a single critical scale. Instead of showing sufficient contribution for *every* pair, we first focus on certain nets of the space at appropriate scales, showing that this suffices to prove the desired lower bound on all pairs. The bulk of the proof consists of arguing that not only the net pairs will have sufficient contribution, but that this will happen with high probability (depending on ε), and with very few dependencies on other net points. To show this we use, as in [1], a localization property of the "padding event" of the random partitions: this event is stochastically independent of the "far away" structure of the partition. The ball expected to be padded is very small (which causes additional distortion), but on the other hand the padding probability is high. The fact that the radius of our padded ball is small (as a function of ε) forces us to define the original distance scales to be also be a function of ε. Finally, to argue that the desired lower bound happens for all pairs with positive probability, even though the number of dimensions at our disposal is small, we use the Lovász Local Lemma.

Acknowledgements. We are grateful to Tim Austin and Bruce Kleiner for helpful discussions.

2. Preliminaries

Due to (1) it suffices to prove Theorem 1.2 when X is finite, provided that the resulting distortion $D(K,\varepsilon)$ and dimension $N(K)$ do not depend on $|X|$. We will therefore assume from now on that X is finite. This assumption is actually not necessary for our argument, but it serves the role of allowing us to ignore measurability issues that might arise in the random partitioning arguments.

For a partition P of X and $x \in X$ let $P(x) \in P$ be the set in P to which x belongs. For $s > 0$ the partition P is called s-bounded if the diameter of $P(x)$ is at most s for all $x \in X$.

There is a canonical way to obtain partitions from balls. Given $x_1,\ldots,x_n\in X$ and $r_1,\ldots,r_n\in(0,\infty)$, define a partition $P^{x_1,\ldots,x_n}_{r_1,\ldots,r_n}$ of $\bigcup_{j=1}^n B(x_j,r_j)$ by

$$P^{x_1,\ldots,x_n}_{r_1,\ldots,r_n}\stackrel{\mathrm{def}}{=}\{B(x_1,r_1)\}\bigcup\left\{B(x_j,r_j)\smallsetminus\bigcup_{i=1}^{j-1}B(x_i,r_i)\right\}_{j=2}^{n}\smallsetminus\{\emptyset\}. \tag{5}$$

In particular, given $s>0$ the partition $P^{x_1,\ldots,x_n}_{r_1,\ldots,r_n}$ is an s-bounded partition of X whenever $\{x_1,\ldots,x_n\}$ is an $s/4$-net of X and $r_1,\ldots,r_n\in[s/4,s/2]$.

As in [1], we will use random partitions of the form $P^{x_1,\ldots,x_n}_{r_1,\ldots,r_n}$, where the radii $r_1,\ldots,r_n$ are appropriately chosen random variables. We present the proofs of the necessary properties of these partitions below, even though they follow from [1]. We do so since the argument of [1] is carried out in much greater generality because in [1] these methods are used for other purposes for which more general constructions are needed. Our argument below is simpler than the proof in [1] both because it deals with the special case that we need, but also because the proof here is different from [1] (relying, of course, on the same ideas).

Lemma 2.1. *Fix $x\in X$. For $s>0$ and $K>1$ let R be a random variable with the following density*

$$\phi_s(r)\stackrel{\mathrm{def}}{=}\frac{16K^8\log K}{s(K^4-1)}K^{-16r/s}\mathbf{1}_{[s/4,s/2]}(r). \tag{6}$$

Then for every $\beta>0$ and every $y\in X$ we have

$$\begin{aligned}\mathbb{P}&\left[B(x,R)\cap B(y,\beta s)\notin\left\{\emptyset,B(y,\beta s)\right\}\right]\\&\qquad\leqslant\left(1-K^{-32\beta}\right)\left(\mathbb{P}\left[B(y,\beta s)\cap B(x,R)\neq\emptyset\right]+\frac{1}{K^4-1}\right).\end{aligned} \tag{7}$$

Proof. Fix $x,y\in X$ and define

$$a\stackrel{\mathrm{def}}{=}\min_{z\in B(y,\beta s)}d(x,z)\quad\text{and}\quad b\stackrel{\mathrm{def}}{=}\max_{z\in B(y,\beta s)}d(x,z).$$

By the triangle inequality,

$$b-a\leqslant 2\beta s. \tag{8}$$

Hence,

$$\begin{aligned}\mathbb{P}&\left[B(x,R)\cap B(y,\beta s)\notin\left\{\emptyset,B(y,\beta s)\right\}\right]=\int_{\max\{a,s/4\}}^{b}\phi_s(r)dr\\&\leqslant\frac{K^8}{K^4-1}\left(K^{-16\max\{a,s/4\}/s}-K^{-16b/s}\right)\stackrel{(8)}{\leqslant}\frac{K^8}{K^4-1}K^{-16\max\{a,s/4\}/s}\left(1-K^{-32\beta}\right).\end{aligned} \tag{9}$$

Similarly,

$$\mathbb{P}\left[B(y,\beta s)\cap B(x,R)\neq\emptyset\right]=\int_{\max\{a,s/4\}}^{s/2}\phi_s(r)dr=\frac{K^8}{K^4-1}\left(K^{-16\max\{a,s/4\}/s}-K^{-8}\right). \tag{10}$$

The desired inequality (7) now follows from (9) and (10). $\square$

Lemma 2.2. *Fix $s>0$, $K\geqslant 2$. Assume that (X,d) has doubling constant at most K. Let $\{x_1,\dots,x_n\}\subseteq X$ be an $s/4$-net of X, and let $R_1,\dots,R_n$ be i.i.d. random variables whose distribution is given by (6). Then for every $y\in X$ and $\beta\in(0,1/40)$ we have,*

$$\mathbb{P}\left[B(y,\beta s)\subseteq P^{x_1,\dots,x_n}_{R_1,\dots,R_n}(y)\right]\geqslant K^{-64\beta}.$$

Proof. For every $j\in\{1,\dots,n\}$ consider the following event:

$$A_j\stackrel{\mathrm{def}}{=}\left(\bigcap_{i=1}^{j-1}\{B(x_i,R_i)\cap B(y,\beta s)=\emptyset\}\right)\bigcap\left\{B(x_j,R_j)\cap B(y,\beta s)\notin\{\emptyset,B(y,\beta s)\}\right\}.$$

For A_j to occur we need in particular to have $B(x_j,R_j)\cap B(y,\beta s)\neq\emptyset$. Since $R_j,\beta s\leqslant s/2$, this implies that $j\in J_y$, where

$$J_y\stackrel{\mathrm{def}}{=}\{j\in\{1,\dots,n\}:\ x_j\in B(y,s)\}.$$

We can cover $B(y,s)$ by at most K^3 balls of radius $s/8$. Since $\{x_1,\dots,x_n\}$ is an $s/4$-net, each of these balls can contain at most one of the x_i. Thus

$$|J_y|\leqslant K^3. \tag{11}$$

We must have $B(x_j,R_j)\cap B(y,\beta s)\neq\emptyset$ for some $j\in J_y$, and therefore using the independence of $R_1,\dots R_n$ we see that,

$$\begin{aligned}1&=\sum_{j\in J_y}\mathbb{P}\left[B(x_j,R_j)\cap B(y,\beta s)\neq\emptyset\ \wedge\ B(x_i,R_i)\cap B(y,\beta s)=\emptyset\ \forall i<j\right]\\ &=\sum_{j\in J_y}\left(\prod_{i=1}^{j-1}\mathbb{P}\left[B(x_i,R_i)\cap B(y,\beta s)=\emptyset\right]\right)\mathbb{P}\left[B(x_j,R_j)\cap B(y,\beta s)\neq\emptyset\right].\end{aligned}\tag{12}$$

Now, by the definition of the partition $P^{x_1,\dots,x_n}_{R_1,\dots,R_n}$ we have

$$\left\{B(y,\beta s)\nsubseteq P^{x_1,\dots,x_n}_{R_1,\dots,R_n}(x)\right\}=\bigcup_{j\in J_y}A_j.$$

Thus, using the independence of $R_1,\dots R_n$ once more,

$$\begin{aligned}1-\mathbb{P}\left[B(y,\beta s)\subseteq P^{x_1,\dots,x_n}_{R_1,\dots,R_n}(x)\right]&\leqslant\sum_{j\in J_y}\mathbb{P}[A_j]\\ &=\sum_{j\in J_y}\left(\prod_{i=1}^{j-1}\mathbb{P}\left[B(x_i,R_i)\cap B(y,\beta s)=\emptyset\right]\right)\mathbb{P}\Big[B(x_j,R_j)\cap B(y,\beta s)\notin\{\emptyset,B(y,\beta s)\}\Big]\\ &\stackrel{(7)\wedge(12)}{\leqslant}\left(1-K^{-32\beta}\right)+\left(1-K^{-32\beta}\right)\frac{|J_y|}{K^4-1}\\ &\stackrel{(11)}{\leqslant}\left(1-K^{-32\beta}\right)\left(1+\frac{K^3}{K^4-1}\right)\\ &\leqslant 1-K^{-64\beta},\end{aligned}\tag{13}$$

Where in (13) we used the fact that since $K\geqslant 2$ and $\beta<1/40$, we have $\frac{K^3}{K^4-1}\leqslant K^{-32\beta}$. Indeed, this is equivalent to $32\beta\leqslant\frac{\log(K-K^{-3})}{\log K}$. But, the function $K\mapsto\frac{\log(K-K^{-3})}{\log K}$ is increasing on $(1,\infty)$, since its derivative is $\frac{4}{(K^5-K)\log K}+\frac{\log(K^4/(K^4-1))}{K(\log K)^2}\geqslant 0$. Thus it suffices to check that $32\beta\leqslant\frac{\log(2-2^{-3})}{\log 2}$, which is true since $\beta<1/40$. $\square$

3. The random embedding

Fix $\varepsilon\in(0,1/2)$, $\theta\in(0,1)$ and $K\geqslant 2$. Write $K=e^\gamma$, and define

$$N\stackrel{\mathrm{def}}{=}\left\lceil\frac{c\gamma}{\theta}\right\rceil=\left\lceil\frac{c\log K}{\theta}\right\rceil, \tag{14}$$

where $c>0$ is a universal constant that will be determined later. It will also be convenient to write

$$\tau\stackrel{\mathrm{def}}{=}\frac{\varepsilon^\theta}{32\gamma^\theta}. \tag{15}$$

Let (X,d) be a finite metric space whose doubling constant is at most K. By normalization assume that $\mathrm{diam}(X)=1$. For every $i\in\mathbb{N}$ let $\{x_1^i,\ldots,x_{n_i}^i\}$ be a $\frac14\tau^{i/(1-\varepsilon)}$-net of X. For every $i,k\in\mathbb{N}$ and $j\in\{1,\ldots,n_i\}$ let R_{ij}^k be a random variable whose density is ϕ_s, as given in (6), with $s=\tau^{i/(1-\varepsilon)}$. We will also use random variables $\{U_i^k(C):\ i,k\in\mathbb{N},\ C\subseteq X\}$, each of which is uniformly distributed on the interval $[0,1]$ (thus for each $i,k\in\mathbb{N}$ we have $2^{|X|}$ such random variables). Throughout the argument below it is assumed that the random variables

$$\left\{R_{ij}^k:\ i,k\in\mathbb{N},\ j\in\{1,\ldots,n_i\}\right\}\bigcup\left\{U_i^k(C):\ i,k\in\mathbb{N},\ C\subseteq X\right\} \tag{16}$$

are mutually independent and defined on some probability space $(\Omega,\mathbb{P})$.

We will now consider the random partitions

$$P_i^k\stackrel{\mathrm{def}}{=}P_{R_{i1}^k,\ldots,R_{in_i}^k}^{x_1^i,\ldots,x_{n_i}^i}, \tag{17}$$

where $P_{R_{i1}^k,\ldots,R_{in_i}^k}^{x_1,\ldots,x_{n_i}}$ is defined as in (5). For $i\in\mathbb{N}$ and $k\in\{1,\ldots,N\}$ define a random mapping $f_i^k:X\to\mathbb{R}$ by

$$f_i^k(x)\stackrel{\mathrm{def}}{=}U_i^k\left(P_i^k(x)\right)\cdot\min\left\{\tau^i,64\gamma\tau^{-\frac{i\varepsilon}{1-\varepsilon}-1}d\left(x,X\smallsetminus P_i^k(x)\right)\right\}. \tag{18}$$

Finally, we define a random embedding $F:X\to\mathbb{R}^N$ as follows:

$$F(x)=\left(\frac{\sum_{i=1}^\infty f_i^1(x)}{\sqrt{N}},\ldots,\frac{\sum_{i=1}^\infty f_i^N(x)}{\sqrt{N}}\right)\in\mathbb{R}^N. \tag{19}$$

Note that by the definition of f_i^k, the sums appearing in (19) converge geometrically.

Although F is random, it satisfies the desired $(1-\varepsilon)$-Hölder condition deterministically. The randomness will enter when we prove that with positive probability $\|F(x)-F(y)\|_2$ satisfies the desired lower bound for all $x,y\in X$.

Lemma 3.1. *For every $x, y \in X$ we have*

$$\|F(x)-F(y)\|_2 \leqslant \max_{k\in\{1,\ldots,N\}} \sum_{i=1}^{\infty} \left|f_i^k(x)-f_i^k(y)\right| \lesssim \frac{\gamma^{(1+\theta)(1-\varepsilon)}}{\varepsilon^{1+\theta}} d(x,y)^{1-\varepsilon}$$
$$\lesssim \left(\frac{\log K}{\varepsilon}\right)^{1+\theta} d(x,y)^{1-\varepsilon}. \quad (20)$$

Proof. We first claim that for all $i \in \mathbb{N}$ and $k \in \{1,\ldots,N\}$ we have

$$\left|f_i^k(x)-f_i^k(y)\right| \leqslant \min\left\{\tau^i, 64\gamma\tau^{-\frac{i\varepsilon}{1-\varepsilon}-1} d(x,y)\right\}. \quad (21)$$

To verify (21) we may assume without loss of generality that $f_i^k(x) > f_i^k(y)$. If $P_i^k(x) \neq P_i^k(y)$ then,

$$f_i^k(x)-f_i^k(y) \leqslant f_i^k(x) \leqslant \min\left\{\tau^i, 64\gamma\tau^{-\frac{i\varepsilon}{1-\varepsilon}-1} d\left(x, X \smallsetminus P_i^k(x)\right)\right\},$$

which is trivially bounded from above by the right hand side of (21) since $y \in X \smallsetminus P_i^k(x)$. If $P_i^k(x) = P_i^k(y) = C$, then it cannot be the case that $f_i^k(y) = U_i^k(C)\,\tau^i$, since otherwise $f_i^k(x) \leqslant f_i^k(y)$, contrary to our assumption. We therefore necessarily have

$$f_i^k(x)-f_i^k(y) = f_i^k(x) - 64U_i^k(C)\gamma\tau^{-\frac{i\varepsilon}{1-\varepsilon}-1} d(y, X\smallsetminus C)$$
$$\leqslant 64U_i^k(C)\gamma\tau^{-\frac{i\varepsilon}{1-\varepsilon}-1}\left(d(x,X\smallsetminus C)-d(y,X\smallsetminus C)\right) \leqslant 64\gamma\tau^{-\frac{i\varepsilon}{1-\varepsilon}-1} d(x,y). \quad (22)$$

Since, by the definition (18), $f_i^k(x), f_i^k(y) \in [0,\tau^i]$, we also have $f_i^k(x)-f_i^k(y) \leqslant \tau^i$. This, in conjunction with (22), concludes the proof of (21).

The first inequality of (20) is an immediate consequence of the definition (19). As $K = e^\gamma$, the third inequality in (20) is a trivial overestimate. Note also that since for all $z \in X$ we have $\|F(z)\|_2 \leqslant \sum_{i=1}^\infty \tau^i \lesssim \tau$, the bound in the second inequality of (20) holds true if $d(x,y) > \tau^{1+\frac{1}{1-\varepsilon}}/(64\gamma)$. We may therefore assume that $d(x,y) \leqslant \tau^{1+\frac{1}{1-\varepsilon}}/(64\gamma)$. Let $m \in \mathbb{N}$ be the integer satisfying

$$\frac{\tau}{64\gamma}\cdot\tau^{\frac{m+1}{1-\varepsilon}} < d(x,y) \leqslant \frac{\tau}{64\gamma}\cdot\tau^{\frac{m}{1-\varepsilon}}. \quad (23)$$

Then

$$\|F(x)-F(y)\|_2 \leqslant \sqrt{\frac{1}{N}\sum_{k=1}^N\left(\sum_{i=1}^\infty\left|f_i^k(x)-f_i^k(y)\right|\right)^2} \leqslant \max_{k\in\{1,\ldots,N\}}\sum_{i=1}^\infty\left|f_i^k(x)-f_i^k(y)\right|$$
$$\overset{(21)}{\lesssim} \frac{\gamma}{\tau}d(x,y)\sum_{i=1}^m \tau^{-\frac{i\varepsilon}{1-\varepsilon}} + \sum_{i=m+1}^\infty \tau^i \overset{(15)}{\lesssim} \frac{\gamma^{1+\theta}}{\varepsilon^\theta}d(x,y)\sum_{i=1}^m\tau^{-\frac{i\varepsilon}{1-\varepsilon}} + \sum_{i=m+1}^\infty\tau^i. \quad (24)$$

We estimate the two sums in (24) separately (recalling that $0 < \varepsilon, \tau < 1/2$):

$$\sum_{i=1}^m \tau^{-\frac{i\varepsilon}{1-\varepsilon}} = \frac{\tau^{-\frac{m\varepsilon}{1-\varepsilon}}-1}{1-\tau^{\frac{\varepsilon}{1-\varepsilon}}} \lesssim \frac{1}{\varepsilon}\tau^{-\frac{m\varepsilon}{1-\varepsilon}} \overset{(23)}{\lesssim} \frac{\tau^\varepsilon}{\varepsilon\gamma^\varepsilon d(x,y)^\varepsilon} \overset{(15)}{\lesssim} \frac{1}{\varepsilon\gamma^{(1+\theta)\varepsilon}d(x,y)^\varepsilon}. \quad (25)$$

Similarly,

$$\sum_{i=m+1}^\infty \tau^i = \frac{\tau^{m+1}}{1-\tau} \overset{(23)}{\lesssim} \frac{\gamma^{1-\varepsilon}}{\tau^{1-\varepsilon}}d(x,y)^{1-\varepsilon} \overset{(15)}{\lesssim} \frac{\gamma^{(1+\theta)(1-\varepsilon)}}{\varepsilon^\theta}d(x,y)^{1-\varepsilon}. \quad (26)$$

The desired bound (20) now follows from substituting (25) and (26) into (24). $\square$

3.1. **The Hölder lower bound holds with positive probability.** For every $i \in \mathbb{N}$ write

$$\delta_i \stackrel{\mathrm{def}}{=} \tau^{\frac{i+2}{1-\varepsilon}} \left(\frac{4\varepsilon}{c^*\gamma}\right)^{\frac{1}{1-\varepsilon}}, \tag{27}$$

where c^* is the implied universal constant in the final inequality of (20). Let $\mathscr{N}_i$ be a δ_i-net of X.

Consider the following set:

$$M \stackrel{\mathrm{def}}{=} \left\{(i,u,v) \in \mathbb{N} \times \mathscr{N}_i \times \mathscr{N}_i : \ \tau^{\frac{i}{1-\varepsilon}} < d(u,v) \leqslant 3\tau^{\frac{i-1}{1-\varepsilon}}\right\}. \tag{28}$$

For every $(i,u,v) \in M$ define $G(i,u,v) \subseteq \{1,\ldots,N\}$ as follows.

$$G(i,u,v) \stackrel{\mathrm{def}}{=} \left\{k \in \{1,\ldots,N\} : \ \left|\sum_{j=1}^{\infty} f_j^k(u) - \sum_{j=1}^{\infty} f_j^k(v)\right| \geqslant \frac{\tau^{i+1}}{2}\right\}. \tag{29}$$

For every $(i,u,v) \in M$ let $E(i,u,v) \subseteq \Omega$ be the following event:

$$E(i,u,v) \stackrel{\mathrm{def}}{=} \left\{|G(i,u,v)| \geqslant \frac{N}{2}\right\}, \tag{30}$$

and consider the event $E \subseteq \Omega$ given by:

$$E \stackrel{\mathrm{def}}{=} \bigcap_{(i,u,v)\in M} E(i,u,v). \tag{31}$$

The relevance of the event E is explained in the following lemma.

Lemma 3.2. *If the event E occurs then for all $x, y \in X$ we have*

$$\|F(x) - F(y)\|_2 \gtrsim \left(\frac{\varepsilon}{\log K}\right)^{2\theta} d(x,y)^{1-\varepsilon}.$$

Proof. Let i be the integer such that

$$\tau^{\frac{i}{1-\varepsilon}} < d(x,y) \leqslant \tau^{\frac{i-1}{1-\varepsilon}} \tag{32}$$

Since $\mathscr{N}_i$ is a δ_i-net, where δ_i is given in (27), there exist $u, v \in \mathscr{N}_i$ such that

$$\max\{d(u,x), d(v,y)\} \leqslant \tau^{\frac{i+2}{1-\varepsilon}} \left(\frac{4\varepsilon}{c^*\gamma}\right)^{\frac{1}{1-\varepsilon}}. \tag{33}$$

Assume that $k \in G(i,u,v)$. By Lemma 3.1 we have

$$\max\left\{\left|\sum_{j=1}^{\infty} f_j^k(u) - \sum_{j=1}^{\infty} f_j^k(x)\right|, \left|\sum_{j=1}^{\infty} f_j^k(v) - \sum_{j=1}^{\infty} f_j^k(y)\right|\right\} \stackrel{(20)\wedge(33)}{\leqslant} c^* \left(\frac{\gamma}{\varepsilon}\right)^{1+\theta} \frac{4\tau^{i+2}\varepsilon}{c^*\gamma} \stackrel{(15)}{=} \frac{\tau^{i+1}}{8}. \tag{34}$$

Since $k \in G(i,u,v)$ it follows that

$$\left|\sum_{j=1}^{\infty} f_j^k(x) - \sum_{j=1}^{\infty} f_j^k(y)\right| \stackrel{(29)\wedge(34)}{\geqslant} \frac{\tau^{i+1}}{2} - 2\cdot\frac{\tau^{i+1}}{8} = \frac{\tau^{i+1}}{4}. \tag{35}$$

Since we are assuming that the event E occurs, the lower bound (35) holds for at least $N/2$ values of $k \in \{1, \dots, N\}$. Thus, by the definition of F,

$$\|F(x) - F(y)\|_2 \gtrsim \tau^{i+1} \overset{(15)\wedge(32)}{\gtrsim} \left(\frac{\varepsilon}{\gamma}\right)^{2\theta} d(x,y)^{1-\varepsilon}. \qquad \square$$

Due to Lemma 3.1 and Lemma 3.2, Theorem 1.2 will be proven (with the bounds claimed in (2), with $\delta = 3\theta$), once we establish the following lemma.

Lemma 3.3. *We have $\mathbb{P}[E] > 0$, provided c in (14) is a large enough universal constant.*

The key tool used in the proof of Lemma 3.3 is the Lovász Local Lemma [11]. The variant of this lemma that is stated below is not the same as the classical formulation of the Lovász Local Lemma, but it is a consequence of it, as explained in [1], where a more general statement is needed. For more information on the Lovász Local Lemma and some of its striking applications, see for example the survey of Alon [2].

Lemma 3.4 (Lovász Local Lemma)**.** *Fix $q \in (0,1)$ and $d \in \mathbb{N}$. Let $\mathcal{A}_1, \mathcal{A}_2, \dots \mathcal{A}_n$ be measurable sets in some probability space $(\Omega, \mathbb{P})$. Let $G = (V, E_G)$ be a graph on the vertex set $V = \{\mathcal{A}_1, \mathcal{A}_2, \dots \mathcal{A}_n\}$ with maximal degree d. Let $\rho : \{\mathcal{A}_1, \mathcal{A}_2, \dots \mathcal{A}_n\} \to \mathbb{N}$ be a mapping that satisfies the condition*

$$\{\mathcal{A}_i, \mathcal{A}_j\} \in E_G \implies \rho(\mathcal{A}_i) = \rho(\mathcal{A}_j).$$

Assume that for any $i \in \{1, \dots, n\}$ we have

$$\mathbb{P}\left[\bigcap_{j \in Q} (\Omega \smallsetminus \mathcal{A}_j) \cap \mathcal{A}_i\right] \leqslant q \mathbb{P}\left[\bigcap_{j \in Q} (\Omega \smallsetminus \mathcal{A}_j)\right]$$

for all

$$Q \subseteq \{j \in \{1, \dots, n\} : \ \{\mathcal{A}_i, \mathcal{A}_j\} \notin E_G \ \wedge \ \rho(\mathcal{A}_i) \geqslant \rho(\mathcal{A}_j)\}.$$

Assume also that

$$eq(d+1) \leqslant 1. \tag{36}$$

Then

$$\mathbb{P}\left[\bigcap_{i=1}^{n} (\Omega \smallsetminus \mathcal{A}_i)\right] > 0$$

To use Lemma 3.4 we proceed as follows. For $(i, u, v) \in M$ consider the following random subset of $\{1, \dots, k\}$:

$$L(i,u,v) \overset{\text{def}}{=} \left\{ k \in \{1, \dots, N\} : \ \left| \sum_{j=1}^{i} f_j^k(u) - \sum_{j=1}^{i} f_j^k(v) \right| \geqslant 2\tau^{i+1} \right\}. \tag{37}$$

For $(i, u, v) \in M$ and $k \in \{1, \dots, N\}$ define the following event:

$$S(i,u,v,k) \overset{\text{def}}{=} \{k \in L(i,u,v)\}. \tag{38}$$

Finally, we also define the following event for all $(i, u, v) \in M$:

$$T(i,u,v) \overset{\text{def}}{=} \left\{ |L(i,u,v)| \geqslant \frac{N}{2} \right\}. \tag{39}$$

Lemma 3.5. *For all $(i, u, v) \in M$ we have $T(i, u, v) \subseteq E(i, u, v)$.*

Proof. Using (21) we see that for all $k \in \{1, \ldots, N\}$,

$$\left| \sum_{j=i+1}^{\infty} f_j^k(u) - \sum_{j=i+1}^{\infty} f_j^k(v) \right| \leqslant \sum_{j=i+1}^{\infty} \tau^j \leqslant \frac{3}{2}\tau^{i+1}. \tag{40}$$

Hence, if $k \in L(i, u, v)$ then

$$\left| \sum_{j=1}^{\infty} f_j^k(u) - \sum_{j=1}^{\infty} f_j^k(v) \right| \geqslant \left| \sum_{j=1}^{i} f_j^k(u) - \sum_{j=1}^{i} f_j^k(v) \right| - \left| \sum_{j=i+1}^{\infty} f_j^k(u) - \sum_{j=i+1}^{\infty} f_j^k(v) \right| \overset{(37)\wedge(40)}{\geqslant} \frac{\tau^{i+1}}{2}.$$

This means that $L(i, u, v) \subseteq G(i, u, v)$, and hence $|L(i, u, v)| \geqslant \frac{N}{2} \implies |G(i, u, v)| \geqslant \frac{N}{2}$. $\square$

Before proceeding with the proof of Lemma 3.3, it is beneficial for us to introduce some notation related to the random partitions that are used in the definition of the embedding F. For $i \in \mathbb{N}$ and $k \in \{1, \ldots, N\}$ the partition P_i^k was defined in (17), where $\{x_1^i, \ldots, x_{n_i}^i\}$ is a fixed $\frac{1}{4}\tau^{i/(1-\varepsilon)}$-net of X, and $R_{i1}^k, \ldots, R_{in_i}^k$ are i.i.d. random variables whose density is ϕ_s as given in (6), with $s = \tau^{i/(1-\varepsilon)}$. For every $y \in X$ define

$$J(i, y) \overset{\text{def}}{=} \left\{ j \in \{1, \ldots, n_i\} : \ d\left(y, x_j^i\right) \leqslant 2\tau^{\frac{i}{1-\varepsilon}} \right\}. \tag{41}$$

We will consider the following random variable

$$j(i, k, y) \overset{\text{def}}{=} \min \left\{ j \in J(i, y) : y \in B\left(x_j^i, R_{ij}^k\right) \right\}. \tag{42}$$

To see that $j(i, k, y)$ is well-defined, note that since $\{x_1^i, \ldots, x_{n_i}^i\}$ is a $\frac{1}{4}\tau^{i/(1-\varepsilon)}$-net of X and $R_{i1}^k, \ldots, R_{in_i}^k \geqslant \frac{1}{4}\tau^{i/(1-\varepsilon)}$, there must be some $j \in \{1, \ldots, N\}$ for which $y \in B\left(x_j^i, R_{ij}^k\right)$, and since $R_{ij}^k \leqslant \frac{1}{2}\tau^{i/(1-\varepsilon)}$ necessarily $j \in J(i, y)$.

¿From the definition (5) we see that

$$P_i^k(y) = B\left(x_{j(i,k,y)}^i, R_{ij(i,k,y)}^k\right) \setminus \bigcup_{\ell=1}^{j(i,k,y)-1} B\left(x_\ell^i, R_{i\ell}^k\right). \tag{43}$$

But note that if there exists $z \in B\left(x_\ell^i, R_{i\ell}^k\right) \cap B\left(x_{j(i,k,y)}^i, R_{ij(i,k,y)}^k\right)$ then

$$d\left(x_\ell^i, y\right) \leqslant d\left(x_\ell^i, z\right) + d\left(z, x_{j(i,k,y)}^i\right) + d\left(x_{j(i,k,y)}^i, y\right) \leqslant R_{i\ell}^k + 2R_{ij(i,k,y)}^k \leqslant 2\tau^{\frac{i}{1-\varepsilon}},$$

implying that $\ell \in J(i, y)$. It follows from this that (43) can be rewritten as follows:

$$P_i^k(y) = B\left(x_{j(i,k,y)}^i, R_{ij(i,k,y)}^k\right) \setminus \bigcup_{\ell \in J(i,y) \cap \{1, \ldots, j(i,k,y)-1\}} B\left(x_\ell^i, R_{i\ell}^k\right). \tag{44}$$

To continue with our plan to use Lemma 3.4, we define a graph $H = (V, E_H)$, where $V \overset{\text{def}}{=} \{T(i, u, v) : \ (i, u, v) \in M\}$, and a mapping $\rho : V \to \mathbb{N}$, as follows.

$$\{T(i, u, v), T(i', u', v')\} \in E_H \iff i = i' \ \wedge \ d(\{u, v\}, \{u', v'\}) \leqslant 4\tau^{\frac{i}{1-\varepsilon}}, \tag{45}$$

$$\rho(T(i, u, v)) = i. \tag{46}$$

Lemma 3.6. *The maximal degree of H is at most $K^{c^{**}(\log \log K + \log(1/\varepsilon))}$, where $c^{**} \in (0, \infty)$ is a universal constant.*

Proof. Given $(i,u,v) \in M$, we need to bound the number of $(i,u',v') \in M$ satisfying $d(\{u,v\},\{u',v'\}) \leqslant 4\tau^{i/(1-\varepsilon)}$. We may assume that $d(u,u') \leqslant 4\tau^{i/(1-\varepsilon)}$. Recall that from the definition of M in (28) we know that $d(u,v), d(u',v') \leqslant 3\tau^{(i-1)/(1-\varepsilon)}$. Hence the points v,u',v' are all in the ball B of radius $r = 4\tau^{(i-1)/(1-\varepsilon)}$ centered at u, implying that the number of (i,u',v') as above is at most $|B \cap \mathscr{N}_i|^2$. Since (X,d) is K-doubling, B can be covered by at most $K^{1+\log_2(2r/\delta_i)}$ balls of radius $\delta_i/2$, each of which contains at most one point from the δ_i-net $\mathscr{N}_i$ (recall (27) for the definition of δ_i). Hence, the maximal degree of H is at most

$$K^{4+2\log_2(r/\delta_i)} = K^{O(\log\log K + \log(1/\varepsilon))}.$$

□

Lemma 3.7. *For every $(i,u,v) \in M$ and for every*

$$Q \subseteq \Big\{(i',u',v') \in M :\ i \geqslant i' \ \wedge\ \{T(i,u,v), T(i',u',v')\} \notin E_H\Big\}, \tag{47}$$

we have

$$\mathbb{P}\left[\bigcap_{(i',u',v')\in Q} T(i',v',u') \smallsetminus T(i,u,v)\right] \leqslant \left(\frac{\varepsilon}{\log K}\right)^{\theta N/2} \mathbb{P}\left[\bigcap_{(i',u',v')\in Q} T(i',v',u')\right]. \tag{48}$$

Proof. Denote

$$W \stackrel{\text{def}}{=} \bigcap_{(i',u',v')\in Q} T(i',v',u').$$

Consider the following subsets $\mathscr{X}, \mathscr{Y}$ of the random variables given in (16):

$$\mathscr{X} \stackrel{\text{def}}{=} \Big\{R^k_{i'j} :\ i' \in \{1,\dots,i-1\},\ j \in \{1,\dots,n_{i'}\},\ k \in \{1,\dots,N\}\Big\}$$
$$\bigcup \Big\{U^k_{i'}(C) :\ i' \in \{1,\dots,i-1\},\ k \in \{1,\dots,N\},\ C \subseteq X\Big\},$$

$$\mathscr{Y} \stackrel{\text{def}}{=} \Big\{R^k_{ij} :\ j \in \{1,\dots,n_i\} \smallsetminus J(i,u),\ k \in \{1,\dots,N\}\Big\}$$
$$\bigcup \Big\{U^k_i(C) :\ k \in \{1,\dots,N\},\ C \subseteq X \smallsetminus B\left(u, 2\tau^{i/(1-\varepsilon)}\right)\Big\}.$$

The event W depends only on the variables $\mathscr{X} \cup \mathscr{Y}$. Indeed, for $(i',u',v') \in Q$ with $i' < i$ it follows from the definitions (39),(37),(18) that the event $T(i',u',v')$ depends only on the variables $\mathscr{X}$. If $(i,u',v') \in Q$ and $\{T(i,u,v), T(i,u',v')\} \notin E_H$ then it follows from (45) that $d(u',u), d(v',u) > 4\tau^{i/(1-\varepsilon)}$. Since the diameter of $P^k_i(u'), P^k_i(v')$ is at most $\tau^{i/(1-\varepsilon)}$, it follows that $P^k_i(u'), P^k_i(v') \subseteq X \smallsetminus B\left(u, 2\tau^{i/(1-\varepsilon)}\right)$, and hence $U^k_i(P^k_i(u')), U^k_i(P^k_i(v')) \in \mathscr{Y}$. Similarly, we know that $J(i,u) \cap J(i,u') = \emptyset$ and $J(i,u) \cap J(i,v') = \emptyset$, and hence from the identity (44) we know that $P^k_i(u'), P^k_i(v')$ depend only on the variables $\mathscr{Y}$. These observations, combined with the definition (18), imply that $f^k_i(u'), f^k_i(v')$ depend only on the variables $\mathscr{Y}$, and from the definitions (39),(37) we conclude that the event $T(i,u',v')$ depends only on the variables $\mathscr{X} \cup \mathscr{Y}$, as required.

Recalling the definitions (37), (38), (39), it follows from the above argument that

$$\mathbb{P}[W \cap T(i,u,v)] = \int_W \mathbb{P}\left[\left.\sum_{k=1}^N \mathbf{1}_{S(i,u,v,k)} \geqslant \frac{N}{2}\ \right|\ \mathscr{X} \cup \mathscr{Y}\right] d\mathbb{P}. \tag{49}$$

To estimate the right hand side of (49), for each $k \in \{1, \dots, n\}$ consider the event

$$Z_k \stackrel{\mathrm{def}}{=} S(i,u,v,k) \cap \left\{ B\left(u, \frac{\tau}{64\gamma}\tau^{\frac{i}{1-\varepsilon}}\right) \subseteq P_i^k(u) \right\}. \tag{50}$$

From (49) we then see that

$$\mathbb{P}\left[W \cap T(i,u,v)\right] \geqslant \int_W \mathbb{P}\left[\sum_{k=1}^N \mathbf{1}_{Z_k} \geqslant \frac{N}{2} \middle| \mathscr{X} \cup \mathscr{Y} \right] d\mathbb{P}. \tag{51}$$

An application of Lemma 2.2 with $\beta = \frac{\tau}{64\gamma}$ yields the estimate

$$\mathbb{P}\left[B\left(u, \frac{\tau}{64\gamma}\tau^{\frac{i}{1-\varepsilon}}\right) \subseteq P_i^k(u) \right] \geqslant K^{-\tau/\gamma} = e^{-\tau}. \tag{52}$$

Moreover, it follows from the definition (18) that if $B\left(u, \frac{\tau}{64\gamma}\tau^{\frac{i}{1-\varepsilon}}\right) \subseteq P_i^k(u)$ then we have $f_i^k(u) = U_i^k\left(P_i^k(u)\right)\tau^i$. Hence, recalling the definition (38), we see that

$$\begin{aligned} Z_k = &\left\{ B\left(u, \frac{\tau}{64\gamma}\tau^{\frac{i}{1-\varepsilon}}\right) \subseteq P_i^k(u) \right\} \\ &\bigcap \left\{ \left| \sum_{j=1}^{i-1} \left(f_j^k(u) - f_j^k(v)\right) + U_i^k\left(P_i^k(u)\right)\tau^i - f_i^k(v) \right| \geqslant 2\tau^{i+1} \right\}. \end{aligned} \tag{53}$$

¿From the identity (44) we see that the event $\left\{ B\left(u, \frac{\tau}{64\gamma}\tau^{\frac{i}{1-\varepsilon}}\right) \subseteq P_i^k(u) \right\}$ is independent of $\mathscr{X} \cup \mathscr{Y}$. Thus, denoting $a \stackrel{\mathrm{def}}{=} \tau^{-i}\sum_{j=1}^{i-1}\left(f_j^k(u) - f_j^k(v)\right) - \tau^{-i} f_i^k(v)$, we have

$$\begin{aligned} p &\stackrel{\mathrm{def}}{=} \mathbb{P}\left[Z_k \middle| \mathscr{X} \cup \mathscr{Y} \right] && (54) \\ &\stackrel{(53)}{=} \mathbb{P}\left[B\left(u, \frac{\tau}{64\gamma}\tau^{\frac{i}{1-\varepsilon}}\right) \subseteq P_i^k(u) \right] \cdot \mathbb{P}\left[U_i^k\left(P_i^k(u)\right) \notin (a - 2\tau, a + 2\tau) \middle| \mathscr{X} \cup \mathscr{Y} \right] \\ &\stackrel{(52)}{\geqslant} e^{-\tau}(1 - 4\tau) && (55) \\ &\geqslant 1 - 5\tau, && (56) \end{aligned}$$

where in (55) we used the fact that since $d(u,v) > \tau^{i/(1-\varepsilon)}$ (by the definition (28) of M), and $\mathrm{diam}(P_i^k(u)) \leqslant \tau^{i/(1-\varepsilon)}$, we have $P_i^k(u) \neq P_i^k(v)$, and therefore the random variable $U_i^k(P_i^k(u))$, which is uniformly distributed on $[0,1]$, is independent of a and $\mathscr{X} \cup \mathscr{Y}$.

Since after fixing the values of $\mathscr{X} \cup \mathscr{Y}$, the events $Z_1, \dots, Z_N$ are independent, the Chernoff bound (see [3, Thm. A.1.12]) implies that

$$\begin{aligned} \mathbb{P}\left[\sum_{k=1}^N \mathbf{1}_{Z_k} \geqslant \frac{N}{2} \middle| \mathscr{X} \cup \mathscr{Y} \right] &= 1 - \mathbb{P}\left[\sum_{k=1}^N \mathbf{1}_{\Omega \setminus Z_k} > \frac{N}{2} \middle| \mathscr{X} \cup \mathscr{Y} \right] \\ &\stackrel{(54)}{\geqslant} 1 - \left(e^{p - \frac{1}{2}} \sqrt{2(1-p)} \right)^N \stackrel{(56)}{\geqslant} 1 - (10e\tau)^{N/2}. \end{aligned} \tag{57}$$

Substituting (57) into (51) shows that

$$\mathbb{P}\left[W \cap T(i,u,v)\right] \geqslant \left(1-(30\tau)^{N/2}\right)\mathbb{P}[W] \overset{(15)}{\geqslant} \left(1-\left(\frac{\varepsilon}{\log K}\right)^{\theta N/2}\right)\mathbb{P}[W],$$

which is the same statement as (48). $\square$

Proof of Lemma 3.3. By Lemma 3.5 we have

$$\mathbb{P}[E] \geqslant \mathbb{P}\left[\bigcap_{(i,u,v)\in M} T(i,u,v)\right].$$

Hence, due to Lemma 3.6 and Lemma 3.7, Lemma 3.3 will follow from Lemma 3.4 if the condition corresponding to (36) holds, i.e.,

$$e\left(\frac{\varepsilon}{\log K}\right)^{\theta N/2}\left(K^{O(\log\log K+\log(1/\varepsilon))}+1\right) \leqslant 1.$$

This holds true provided the constant c in the definition (14) of N is large enough. $\square$

4. Snowflakes of the Heisenberg group

As promised in the introduction, we first argue that a positive answer to the qualitative version of the Lang-Plaut question, as appearing in Question 1, implies its quantitative variant, i.e., that for every $K>0$ there is $N=N(K)\in\mathbb{N}$ and $D=D(K)\in(1,\infty)$ such that if $X\subseteq \ell_2$ has doubling constant K then $c_{\mathbb{R}^N}(X)\leqslant D$. Indeed, if not then there would be some $K>0$ and a sequence $\{X_n\}_{n=1}^\infty$ of subsets of ℓ_2 with doubling constant K and satisfying $c_{\mathbb{R}^n}(X_n)>n$. By (1) there are finite subsets $F_n\subseteq X_n$ with $c_{\mathbb{R}^n}(F_n)>n$, and by translation and rescaling we may assume that $0\in F_n$ and that F_n is contained in the ball of ℓ_2 centered at 0 of radius 1. Let $Y\subseteq \ell_2\times\mathbb{R}$ be given by $Y=\bigcup_{n=1}^\infty F_n\times\{4^n\}$. One checks that Y has doubling constant $O(K)$, and clearly all the F_n embed into Y isometrically. By the assumed positive answer to the Lang-Plaut problem it follows that $c_{\mathbb{R}^N}(F_n)\leqslant D$ for some $N\in\mathbb{N}$ and $D\in(1,\infty)$, a contradicting the fact that $c_{\mathbb{R}^N}(F_n)>n$ for $n\geqslant N$.

Proof of Theorem 1.3. For $\theta>0$ define $\delta_\theta:\mathbb{H}_n\to\mathbb{H}_n$ by $\delta_\theta(z,t)=(\theta z,\theta^2 t)$. Note that for every measurable $A\subseteq\mathbb{H}_n=\mathbb{C}^n\times\mathbb{R}$ we have $\mathrm{vol}(\delta_\theta(A))=\theta^{2n+2}\mathrm{vol}(A)$. For $p\in[1,2)$ and $(z,t)\in\mathbb{H}_n$ define

$$M_p(z,t)\stackrel{\mathrm{def}}{=}\sqrt[4]{|z|^4+t^2}\left(\cos\left(\frac{p}{2}\arccos\left(\frac{|z|^2}{\sqrt{|z|^4+t^2}}\right)\right)\right)^{1/p}.$$

It was shown in [22] that $M_p(xy^{-1})\leqslant M_p(x)+M_p(y)$ for all $x,y\in\mathbb{H}_n$. Therefore $d_{M_p}(x,y)=M_p(y^{-1}x)$ is a left-invariant metric on $\mathbb{H}_n$. It was also shown in [22] that $\sqrt{1-\frac{p}{2}}N_0(x)\leqslant M_p(x)\leqslant N_0(x)$ for all $x\in\mathbb{H}_n$, and there exists $f:\mathbb{H}_n\to\ell_2$ satisfying $\|f(x)-f(y)\|=d_{M_p}(x,y)^{p/2}$ for all $x,y\in\mathbb{H}_n$. Setting $p=2(1-\varepsilon)$, we see that for all distinct $x,y\in\mathbb{H}_n$ we have

$$\frac{\|f(x)-f(y)\|}{d_{N_0}(x,y)^{1-\varepsilon}}=\left(\frac{d_{M_p}(x,y)}{d_{N_0}(x,y)}\right)^{p/2}\in\left[\varepsilon^{(1-\varepsilon)/2},1\right]\subseteq[\sqrt{\varepsilon},1].$$

For $x \in \mathbb{H}_n$ and $r > 0$ denote $B_p(x,r) = \{y \in \mathbb{H}_n : \ d_{M_p}(x,y)^{p/2} \leqslant r\}$. Note that $B_p(0,r) = \delta_{2^{2/p}}(B_p(0,r/2))$, since for every $\theta > 0$ and $x, y \in \mathbb{H}_n$ we have $d_{M_p}(\delta_\theta(x), \delta_\theta(y)) = \theta d_{M_p}(x,y)$. Hence, by left-invariance of $d_{M_p}^{p/2}$ and the Lebesgue measure $\mathrm{vol}(\cdot)$, for all $x \in \mathbb{H}_n$ and $r > 0$ we have $\mathrm{vol}(B_p(x,r)) = 2^{4(n+1)/p}\mathrm{vol}(B_p(x,r/2))$. This implies that $\left(\mathbb{H}_n, d_{M_p}^{p/2}\right) = \left(\mathbb{H}_n, d_{M_p}^{1-\varepsilon}\right)$, and hence also its isometric copy $f(\mathbb{H}_n) \subseteq \ell_2$, has doubling constant $2^{8(n+1)/p} \leqslant 2^{16(n+1)}$. $\square$

We end with the proof of the distortion lower bound (4). Assume that $f : \mathbb{H}_1 \to \ell_2$ satisfies

$$d_{N_0}(x,y)^{1-\varepsilon} \leqslant \|f(x) - f(y)\| \leqslant D d_{N_0}(x,y)^{1-\varepsilon} \quad \forall\, x, y \in \mathbb{H}_1. \tag{58}$$

Our goal is to prove that $D \gtrsim 1/\sqrt{\varepsilon}$. Denote $a = (1,0) \in \mathbb{H}_1$, $b = (i,0) \in \mathbb{H}_1$ and $c = aba^{-1}b^{-1} = (0,-4)$. Writing

$$B_m = \{(u+iv,t) \in \mathbb{H}_1 : \ u,v,t \in \mathbb{Z} \ \wedge \ N_0(u+iv,t) \leqslant m\},$$

it follows from [5] that there exists a universal constant $C > 0$ such that for all $m \in \mathbb{N}$ we have

$$\sum_{x \in B_m} \sum_{k=1}^{m^2} \frac{\left\|f(xc^k) - f(x)\right\|^2}{k^2} \lesssim \sum_{x \in B_{Cm}} \left(\|f(xa) - f(x)\|^2 + \|f(xb) - f(x)\|^2\right). \tag{59}$$

Note that for all $x \in \mathbb{H}_1$ we have $d_{N_0}(xa,x) = N_0(a) = 1$ and similarly $d_{N_0}(xb,x) = 1$. Moreover, for all $k \in \mathbb{N}$ and $x \in \mathbb{H}_1$ we have $d_{N_0}(xc^k,x) = N_0(c^k) = N_0(0,-4k) = 2\sqrt{k}$. Hence, using (58) and the fact that the cardinality of B_m is bounded above and below by universal multiples of m^4, inequality (59) becomes $D^2 \gtrsim \sum_{k=1}^{m^2} \frac{N_0(c^k)^{2(1-\varepsilon)}}{k^2} \gtrsim \sum_{k=1}^{m^2} \frac{1}{k^{1+\varepsilon}}$. Letting m tend to ∞ we deduce that $D^2 \gtrsim \sum_{k=1}^{\infty} \frac{1}{k^{1+\varepsilon}} \gtrsim \frac{1}{\varepsilon}$, as required. $\square$

References

[1] I. Abraham, Y. Bartal, and O. Neiman. Embedding metric spaces in their intrinsic dimension. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 363–372, New York, 2008. ACM.

[2] N. Alon. Probabilistic proofs of existence of rare events. In *Geometric aspects of functional analysis (1987–88)*, volume 1376 of *Lecture Notes in Math.*, pages 186–201. Springer, Berlin, 1989.

[3] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., Hoboken, NJ, third edition, 2008. With an appendix on the life and work of Paul Erdős.

[4] P. Assouad. Plongements lipschitziens dans $\mathbf{R}^n$. *Bull. Soc. Math. France*, 111(4):429–448, 1983.

[5] T. Austin, A. Naor, and R. Tessera. Sharp quantitative nonembeddability of the Heisenberg group into superreflexive Banach spaces. Preprint available at `http://www.cims.nyu.edu/~naor/homepage%20files/UCHeis9.pdf`, 2010.

[6] Y. Bartal, B. Recht, and L. Schulman. Dimensionality reduction: beyond the Johnson-Lindenstrauss bound. To appear in Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'2011). Preprint available at `http://pages.cs.wisc.edu/~brecht/papers/07.BRS.local-dim-REVISED.pdf`, 2011.

[7] M. Bourdon and H. Pajot. Poincaré inequalities and quasiconformal structure on the boundary of some hyperbolic buildings. *Proc. Amer. Math. Soc.*, 127(8):2315–2324, 1999.

[8] J. Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.

[9] J. Cygan. Subadditivity of homogeneous norms on certain nilpotent Lie groups. *Proc. Amer. Math. Soc.*, 83(1):69–70, 1981.

[10] G. David and T. Toro. Reifenberg flat metric spaces, snowballs, and embeddings. *Math. Ann.*, 315(4):641–710, 1999.

[11] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II*, pages 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10. North-Holland, Amsterdam, 1975.

[12] L.-A. Gottlieb and R. Krauthgamer. A nonlinear approach to dimension reduction. To appear in Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'2011). Preprint available at `http://arxiv.org/abs/0907.5477`, 2011.

[13] A. Gupta, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *44th Symposium on Foundations of Computer Science*, pages 534–543. IEEE Computer Society, 2003.

[14] 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.

[15] J. Heinonen. *Lectures on analysis on metric spaces.* Universitext. Springer-Verlag, New York, 2001.

[16] J. Heinonen. *Geometric embeddings of metric spaces*, volume 90 of *Report. University of Jyväskylä Department of Mathematics and Statistics*. University of Jyväskylä, Jyväskylä, 2003.

[17] J.-P. Kahane. Hélices et quasi-hélices. In *Mathematical analysis and applications, Part B*, volume 7 of *Adv. in Math. Suppl. Stud.*, pages 417–433. Academic Press, New York, 1981.

[18] A. Korányi and H. M. Reimann. Quasiconformal mappings on the Heisenberg group. *Invent. Math.*, 80(2):309–338, 1985.

[19] 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.

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

[21] 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.

[22] J. R. Lee and A. Naor. L_p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *47th Annual IEEE Symposium on Foundations of Computer Science*, pages 99–108, 2006. Available at `http://www.cims.nyu.edu/~naor/homepage%20files/L_pHGL.pdf`.

[23] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[24] A. Naor and L. Silberman. Poincaré inequalities, embeddings, and wild groups. Preprint available at `http://www.cims.nyu.edu/~naor/homepage%20files/wilder.pdf`, 2010.

[25] P. Pansu. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math. (2)*, 129(1):1–60, 1989.

[26] I. J. Schoenberg. Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.*, 44(3):522–536, 1938.

[27] S. Semmes. On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞-weights. *Rev. Mat. Iberoamericana*, 12(2):337–410, 1996.

[28] S. Semmes. Bilipschitz embeddings of metric spaces into Euclidean spaces. *Publ. Mat.*, 43(2):571–653, 1999.

[29] M. Talagrand. Approximating a helix in finitely many dimensions. *Ann. Inst. H. Poincaré Probab. Statist.*, 28(3):355–363, 1992.

Courant Institute, New York University, New York NY 10012
E-mail address: `naor@cims.nyu.edu`

Computer Science Department, Princeton University, Princeton NJ 08544-2087
E-mail address: `oneiman@princeton.edu`