Low Dimensional Embeddings of Doubling Metrics

Ofer Neiman *

July 24, 2013

Abstract

We study several embeddings of doubling metrics into low dimensional normed spaces, in particular into ℓ_2 and ℓ_∞. Doubling metrics are a robust class of metric spaces that have low intrinsic dimension, and often occur in applications. Understanding the dimension required for a concise representation of such metrics is a fundamental open problem in the area of metric embedding. Here we show that the n-vertex Laakso graph can be embedded into constant dimensional ℓ_2 with the best possible distortion, which has implications for possible approaches to the above problem.

Since arbitrary doubling metrics require high distortion for embedding into ℓ_2 and even into ℓ_1, we turn to the ℓ_∞ space that enables us to obtain arbitrarily small distortion. We show embeddings of doubling metrics and their "snowflakes" into low dimensional ℓ_∞ space that simplify and extend previous results.

1 Introduction

In this paper we study embeddings of doubling metric spaces into low dimension normed spaces. A metric space (X, d) has doubling constant λ if any ball can be covered by λ balls of half its radius. A family of metrics is called *doubling* if the doubling constant of every member is bounded by some universal constant. The past decade has seen a surge of interest in doubling metrics, mainly because numerous algorithmic tasks are (approximately) tractable in such metrics, e.g. routing in networks, low stretch spanners, nearest neighbor search, approximate distance oracles, traveling salesperson problem [HPM06, CGMZ05, GR08, GK11, BGK12].

Embedding into normed spaces is a very useful paradigm for representing and analyzing data. Since the cost of many data processing tasks depend exponentially on the dimension (the "curse of dimensionality"), it is often crucial to obtain a low dimension in the host space. The doubling constant of the metric captures in some sense the intrinsic dimension of the metric, and the logarithm of the doubling constant is known as the *doubling dimension* [GKL03]. Indeed, there are numerous results on low dimensional embedding of doubling metrics, and in what follows we review some of them. Recall that an embedding of a metric space (X, d) into ℓ_p^D is a map $f : X \to \mathbb{R}^D$, and the distortion of f is defined as

$$\max_{x \neq y \in X} \left\{ \frac{\|f(x) - f(y)\|_p}{d(x, y)} \right\} \cdot \max_{x \neq y \in X} \left\{ \frac{d(x, y)}{\|f(x) - f(y)\|_p} \right\} .$$

Several results only hold for a "snowflake" version of the metric: The $1 - \alpha$ snowflake of (X, d) is the metric $(X, d^{1-\alpha})$ with $0 < \alpha < 1$ (that is, every distance is raised to power $1 - \alpha$).

*Department of Computer Science, Ben-Gurion University of the Negev. E-mail: neimano@cs.bgu.ac.il.

Euclidean Embeddings: Assouad [Ass83] showed that if (X, d) is λ-doubling then $(X, d^{1-\alpha})$ can be embedded into constant dimensional Euclidean space with constant distortion, where the constants depend only on λ and on α. He conjectured that such a result is possible also when $\alpha = 0$ (i.e. the original metric), but this was disproved by Semmes [Sem96]. In the computer science community, [GKL03] gave a comprehensive study on embedding doubling metrics. Among other results, they showed that n-point doubling metric spaces can be embedded with tight distortion $O(\sqrt{\log n})$ into Euclidean space (in contrast with arbitrary metrics that may require $\Omega(\log n)$ distortion [LLR95]). [KLMN05] showed an embedding with optimal dependence on the doubling constant (the lower bound was given by [JLM09]). The "price" paid for obtaining optimal distortion is that the dimension of all these embeddings is at least $\Omega(\log n)$. Following the intuition that the doubling dimension should be related to the dimension of the host space, [ABN08] showed that for any $\epsilon > 0$, λ-doubling metrics can be embedded into $O((\log \lambda)/\epsilon)$ dimensional Euclidean space with distortion $O(\log^{1+\epsilon} n)$. Both [ABN08, CGT10] exhibited a tradeoff between distortion and dimension: as the dimension ranges from $O(\log \log n)$ to $O(\log n)$, the distortion ranges from $O(\log n)$ to $O(\sqrt{\log n})$. However, the following is still open:

Question 1. *Does every doubling metric on n points embeds into $O(1)$ dimensional ℓ_2 space with distortion $O(\sqrt{\log n})$?*

Here we (arguably) show some evidence for a positive answer to this question, by providing an embedding of the metric induced by an n-vertex Laakso graph into constant dimensional Euclidean space with distortion $O(\sqrt{\log n})$. The Laakso graph G_k is a series-parallel graph with 6^k edges, $\Theta(6^k)$ vertices, and its doubling constant is at most 6 (see Section 3 for a definition of the Laakso graph), it was first introduced by [Laa02]. This graph seems difficult for ℓ_2 embedding and low dimensional embeddings. In particular, it is known that the metric induced by the n-vertex Laakso graph requires $n^{\Omega(1/\beta^2)}$ dimensions for a β distortion embedding into ℓ_1 [LMN05] (following the results of [BC05, LN04]). Also, this metric requires distortion at least $\Omega(\sqrt{\log n})$ for any embedding into ℓ_2 [GKL03]. So it seems surprising that allowing distortion $O(\sqrt{\log n})$ the embedding only requires 3 dimensions[1].

Theorem 1. *For any positive integer m, there exists an embedding of the metric induced by G_m into 3 dimensional ℓ_2 space with distortion $O(\sqrt{m})$.*

The proof of Theorem 1 appears in Section 3.

Embedding into ℓ_∞: The distortion of the above results is often undesirably high, in particular for application areas, where it is useful to have arbitrarily low distortion. Obtaining low distortion was shown to be impossible for ℓ_2 by [Sem96, Laa02, GKL03], and for ℓ_1 by [CK10, CKN09, LS11], where for the former the lower bound is a tight $\Omega(\sqrt{\log n})$ and for the latter $\Omega(\sqrt{\log n / \log \log n})$. Another natural candidate space is the ℓ_∞ space. In [GKL03] it was shown that for any $\epsilon > 0$, any doubling metric space (X, d) on n points embeds into $\ell_\infty^{O(\log n)}$ with distortion $1 + \epsilon$. While the explicit proof and the dependence on the parameters ϵ and λ was not specified there, the proof was based on a variation of Bourgain's embedding and an application of the Lovász Local Lemma. In this paper we give a very simple proof of this result that does not require the Local Lemma, and has the best possible dependence on ϵ and the doubling constant λ, up to a constant in the exponent. Another advantage is that our construction only requires building nets, which can be implemented efficiently in near linear time [HPM06]. The result is in fact a simple adaptation of the methods introduced by [HPM06].

Theorem 2. *For any $0 < \epsilon \leq 1$, any finite metric space (X, d) on n points with doubling constant λ embeds into ℓ_∞^D with distortion $1 + \epsilon$ where $D = \lambda^{\log(1/\epsilon)+O(1)} \log n$.*

The proof of Theorem 2 appears in Section 4.

[1]It is quite conceivable that 2 dimensions suffice, we used 3 to simplify the analysis.

Snowflake embeddings: Following the result of Assouad, there were several extensions for the snowflakes of doubling metrics. [GKL03] provided an improved dependence of the distortion and the dimension on the doubling constant λ in Assouad's result. The dependence on α in the dimension was further improved in [ABN08], and finally was completely removed in [NN12] (in the range $0 < \alpha < 1/2$). [HPM06], among other algorithmic results on doubling metrics, showed an embedding of $(X, d^{1/2})$ into ℓ_∞ of dimension $\lambda^{O(\log(1/\epsilon))}$, which is then used for distance labeling. More recently, [GK11] showed a dimension reduction result for a snowflake of Euclidean subsets that are doubling, and [BRS11] obtained similar result. For the ℓ_∞ host, they showed a $1+\epsilon$ distortion embedding for a $1-\alpha$ snowflake with $\lambda^{O(\log(1/\epsilon)+\log\log\lambda)}/(\alpha(1-\alpha))$ dimensions. The proof of [GK11] ingeniously combined many "hammers" such as the Johnson-Lindenstrauss dimension reduction, padded decompositions, a Gaussian transform and smoothing techniques. In this work we improve slightly the result of [GK11] for embedding doubling snowflakes into ℓ_∞, and generalize the embedding result of [HPM06] to arbitrary snowflaking parameter α. Perhaps more importantly, the construction and analysis given here are arguably simpler that those of [GK11], and admit an efficient implementation.

Theorem 3. *For any $0 < \epsilon \le 1/20$, $0 < \alpha < 1$, and any finite metric space (X, d) on n points with doubling constant λ, there exists an embedding of the snowflake $(X, d^{1-\alpha})$ into ℓ_∞^D with distortion $1+\epsilon$ where $D = \lambda^{\log(1/\epsilon)+O(1)}/(\alpha(1-\alpha))$.*

The proof of Theorem 3 appears in Section 5.

Dimension Reduction for Doubling Subsets: Assouad's result (embedding doubling snowflakes into constant dimensional Euclidean space with constant distortion) cannot be extended to arbitrary doubling metrics as mentioned above. One of the major open problems in the area of metric embedding is whether his result can be extended to doubling subsets of Euclidean space. That is,

Question 2. *Does every doubling subset of ℓ_2 embeds into constant dimensional ℓ_2 space with constant distortion?*

This question was raised by [LP01, GKL03], and also referred to in other works such as [ABN08, CGT10, GK11, NN12]. A possible approach for finding a counterexample, mentioned in [NN12], is to use the image under Euclidean embedding of a known "difficult" doubling metric. If it can be shown that a certain n-point doubling metric has the following properties: 1) It has an ℓ_2 embedding with distortion $O(\sqrt{\log n})$ in which its image is doubling, and 2) Any embedding of this metric into constant dimensional ℓ_2 requires $\omega(\sqrt{\log n})$ distortion, then it would provide a negative answer to the above question.

A natural candidate for such a doubling metric, used in [CK10, CKN09] to prove non-embeddability in ℓ_1 of negative type metrics, is the Heisenberg group $\mathbb{H}$ equipped with the Carnot-Carathéodory metric. It was shown in [NN12] that it satisfies the first property. Another possible "difficult" metric is the Laakso graph, however the result stated in Theorem 1 rules out this example. In fact, a positive answer to Question 1 would rule out this approach entirely.

2 Preliminaries

Let (X, d) be a finite metric space, with $|X| = n$. We shall assume w.l.o.g that $d(x, y) \ge 1$ for all $x, y \in X$. The diameter of (X, d) is $\operatorname{diam}(X) = \max_{x,y\in X}\{d(x, y)\}$. A ball around $x \in X$ with radius $r \ge 0$ is defined as $B(x, r) = \{z \in X \mid d(x, z) \le r\}$. The doubling constant of (X, d) is the minimal integer λ such that for all $x \in X$ and $r > 0$, the ball $B(x, 2r)$ can be covered by λ balls of radius r. The doubling dimension of (X, d) is defined as $\dim(X) = \log_2 \lambda$. A family of metric spaces is called doubling if there is a constant K such that every metric in the family has doubling constant at most K. An r-net of (X, d) is a set of points $N \subseteq X$ satisfying: 1) For all $u, v \in N$, $d(u, v) > r$, and 2) $\bigcup_{u\in N} B(u, r) = X$. It is well known that a simple greedy algorithm can provide an r-net.

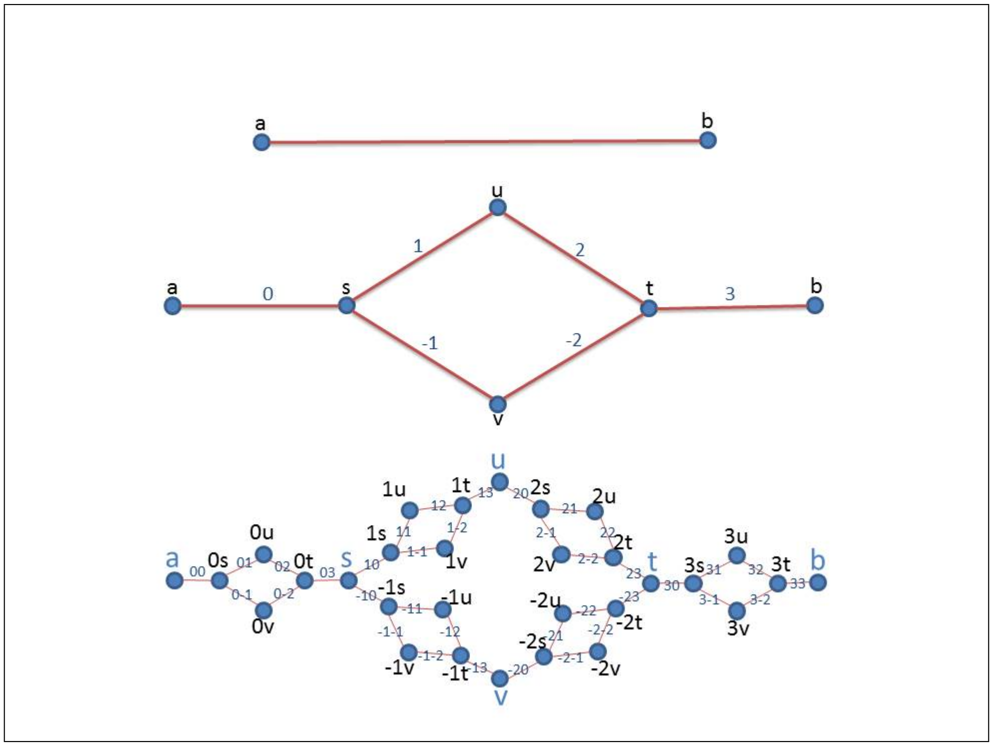

Figure 1: Naming vertices and edges of Laakso graph

3 Low dimensional Embedding of the Laakso Graph

In this section we prove Theorem 1. For integer $k \geq 0$ let G_k be the k-th level Laakso graph, defined as follows: G_0 consist of a single edge, G_k is defined by replacing every edge of G_{k-1} with the graph on six edges and six vertices depicted in Figure 1, such that the vertices a, b correspond to the original endpoints of the edge. The edge lengths in G_k are 4^{-k} for all edges. For a pair of vertices that were edges in G_i, we abuse notation and call them *level i edges*. A level i edge e is a child of a level $i-1$ edge e' if it is one of the six edges that replaced e'. This defines an (partial) inheritance relation on the edges of different levels. Note that any edge at level $k > i$ has a unique level i ancestor.

We label the edges of G_k by a sequence $l \in L^k$, where $L = \{0, 1, -1, 2, -2, 3\}$, such that for $1 \leq i \leq k$, l_i is the position of the level i ancestor of the edge, depicted in Figure 1. The vertices created in level k are labeled by a string in $L^{k-1} \times \{s, t, u, v\}$. In particular, each edge of level $k-1$ creates 4 new vertices, if the label of the edge was $l \in L^{k-1}$, then the new vertices will be labeled by $l \circ s$, $l \circ t$, $l \circ u$, and $l \circ v$ (where for strings w, w', $w \circ w'$ denotes their concatenation).

We write $\|\cdot\|$ for the standard Euclidean norm.

3.1 Construction of the Embedding

Consider the graph G_n, with shortest path metric d, and fix $D = 1/\sqrt{n}$. First define an embedding $g : V(G_n) \to \mathbb{R}$ by $g(x) = d(x, a)$, where a is the left vertex of G_0. We define the embedding $f : V(G_n) \to \mathbb{R}^2$ recursively as follows. In the case $k = 0$ where a, b are the two endpoints of the single edge of level 0, define $f(a) = (0, 0)$, $f(b) = (1, 0)$. Fix some integer $1 \leq k \leq n$. Now, let $\{a, b\}$ be any level $k-1$ edge, let s, t, u, v be the new four vertices created from it in level k. Inductively, f is already defined on both a and b, so let $z = f(b) - f(a)$. Finally let $\bar{z}$ be one of the two unit vectors orthogonal to z in $\mathbb{R}^2$ (chosen arbitrarily). Define

$$f(s) = f(a) + z/4$$
$$f(t) = f(a) + 3z/4$$
$$f(u) = f(a) + z/2 + D4^{-k} \cdot \bar{z}$$
$$f(v) = f(a) + z/2 - D4^{-k} \cdot \bar{z} \,.$$

In some sense the embedding g is just a projection of the graph into the line, and its sole purpose is to provide contribution for the edges. The difficulty in embedding the Laakso graph comes from handling the diagonals (each diagonal is composed of the two vertices whose labels are $p \circ u$ and $p \circ v$ for some $k \geq 0$ and $p \in L^k$). The map f provides sufficient contribution for the diagonals, the price is that we expand slightly the four inner edges (e.g. $\{p \circ s, p \circ u\}$). Intuitively, since f provides only $1/\sqrt{n}$ fraction of the distance between the diagonals, and uses an orthogonal vector to the parent edge's vector, we get that the distance between the images of the edge's endpoints is increased only by a factor of $1/n$. Thus even n levels of recursion will not generate a large expansion.

3.2 Analysis of the Embedding

The first step is to bound the distortion of the edges (of all levels), which yields an upper bound on the expansion of the embedding.

Claim 1. *For any integer $0 \leq k \leq n$ and any level k edge $\{x, y\} \in E(G_k)$,*

$$d(x, y) \leq \|f(x) - f(y)\| \leq 2d(x, y) \,.$$

Proof. We prove by induction on k that if $\{x, y\}$ is level k edge, then

$$4^{-k} \leq \|f(x) - f(y)\| \leq \sqrt{1 + kD^2} \cdot 4^{-k} \,. \tag{1}$$

The base case $k = 0$ is true by definition. For the inductive step, let $\{a, b\}$ be a level $k - 1$ edge with $d(a, b) = 4^{-(k-1)}$, and let $z = f(b) - f(a)$. By the induction hypothesis

$$4^{-(k-1)} \leq \|z\| \leq \sqrt{1 + (k-1)D^2} \cdot 4^{-(k-1)} \,. \tag{2}$$

Consider the six edges created from $\{a, b\}$ in level k that are depicted in Figure 1. First observe that for the edge $\{a, s\}$, by definition $\|f(s) - f(a)\| = \|z\|/4$ so it satisfies (1). The same holds for the edge $\{b, t\}$. Consider now the edge $\{s, u\}$, using that $z, \bar{z}$ are orthogonal suggests the following bound,

$$\begin{aligned} \|f(u) - f(s)\|^2 &= \|z/4 + D4^{-k} \cdot \bar{z}\|^2 \\ &= \|z/4\|^2 + \|D4^{-k} \cdot \bar{z}\|^2 \\ &= \|z\|^2/16 + (D4^{-k})^2 \end{aligned}$$

Using (2) it holds that $\|z\| \geq 4^{-(k-1)}$ and thus $\|f(u) - f(s)\| \geq \|z\|/4 \geq 4^{-k}$. For the upper bound, note that by (2)

$$\|f(u) - f(s)\|^2 \leq (1 + (k-1)D^2) \cdot 4^{-2(k-1)}/16 + D^2 4^{-2k} = (1 + kD^2) \cdot 4^{-2k} \,.$$

The same calculation holds for the edges $\{s, v\}$, $\{u, t\}$ and $\{v, t\}$. This concludes the proof of (1). Using that $k \leq n$, we see that $\sqrt{1 + kD^2} \leq \sqrt{1 + 1} < 2$, proving the claim. ☐

Lemma 2. *For any $x, y \in V(G_n)$,*

$$\|(f \oplus g)(x) - (f \oplus g)(y)\| \leq 3d(x, y) \,.$$

Proof. Let $x = u_0, u_1, \ldots, u_t = y$ be a shortest path in G_n connecting x to y. By the triangle inequality and Claim 1,

$$\|f(x) - f(y)\| \leq \sum_{i=1}^{t} \|f(u_i) - f(u_{i-1})\| \leq 2\sum_{i=1}^{t} d(u_i, u_{i-1}) = 2d(x, y) . \tag{3}$$

Using the triangle inequality it follows that

$$|g(x) - g(y)| = |d(x, a) - d(y, a)| \leq d(x, y) .$$

□

The main effort will be showing that the contraction of $f \oplus g$ is bounded by $O(1/D)$. Observe that the vertices u, v of any basic structures in any level already suffer contraction of $\Theta(1/D)$. If we consider two vertices x, y of distance $d(x, y) \approx 4^{-j}$, then in level j they have different ancestor edges, and at least intuitively they should get a contribution of $D \cdot 4^{-j}$ from the embedding of level j. However, in their final embedding, x, y may "get closer" to each other because we use few dimensions. We first focus on the case where $g(x) = g(y)$, and so the contribution must entirely come from the f embedding. The following lemma shows there is indeed sufficient contribution from the critical level, and the main issue is showing that this contribution does not completely cancel out.

Lemma 3. *Let $x, y \in V(G_n)$ be such that $g(x) = g(y)$, then*

$$\|f(x) - f(y)\| \geq D \cdot d(x, y)/32 .$$

Proof. First observe that since x, y have the same g value they must have been created in the same level, and denote this level by m. Abusing notation, denote by $x, y \in L^{m-1} \times \{s, t, u, v\}$ the labels of the vertices. For any $1 \leq i \leq m$ let $p_i = x_1 \circ \cdots \circ x_{i-1}$ be the label of the level $i - 1$ ancestor edge of x. Observe that since $g(x) = g(y)$ it must be for every $i \in [m - 1]$ that $|x_i| = |y_i|$. First consider the case that $x_i = y_i$ for all $i \in [m]$, then in fact x, y are the u, v vertices created from the edge labeled p_m, and by definition of f, $\|f(x) - f(y)\| = 2D \cdot 4^{-m} = D \cdot d(x, y)$.

Now assume that there is an index $i \in [m - 1]$ such that $x_i \neq y_i$, and let j be the minimal such index. We shall assume w.l.o.g that $x_j = 1$, $y_j = -1$ (the case $x_j = 2$ and $y_j = -2$ is symmetric). Let $s' = p_j \circ s$ and $t' = p_j \circ t$. Let k be the smallest integer satisfying $j < k < m$ and such that at least one of x_k, y_k is different from 0 (or different from 3 if it was the case that $x_j = 2$, $y_j = -2$). If no such k exists put $k = m$. Assume w.l.o.g that $x_k \neq 0$. Roughly speaking, j is the index of the scale in which x, y are separated into different recursive components, however since in the scales i from $j + 1$ to $k - 1$, $x_i = y_i = 0$, both are still close to s' and thus to each other. The final distance between x, y is about 4^{-k}. To see this, note that $d(x, y) = d(x, s') + d(y, s')$, and thus $d(x, y) \geq d(x, s') \geq 4^{-k}$. On the other hand, since for $j < i < k$, $x_i = y_i = 0$, then $d(x, s') \leq d(p_{k-1} \circ s, s') = 4^{-(k-1)}$ and similarly for $d(y, s')$, so $d(x, y) < 4^{-(k-2)}$. It remains to show that $\|f(x) - f(y)\| \geq D \cdot 4^{-k}/2$.

Let ℓ be the line passing through the endpoints of p_j. We will prove that $f(x)$ is at least $D \cdot 4^{-k}/2$ away from any point on the line ℓ. To this end, we prove by induction on m that the Euclidean distance of $f(x)$ from the line ℓ is at least

$$D \cdot (4^{-k} - \sum_{i=k+1}^{m} 4^{-i}) , \tag{4}$$

furthermore, $f(x)$ is on the same side of ℓ as $f(u')$ where $u' = p_j \circ u$.

The base case is when $m = k$. As $x_j = 1$, the end points of the level j ancestor of x are $s' = p_j \circ s$ and $u' = p_j \circ u$. By definition $f(s')$ lies right on the line ℓ connecting the images of the endpoints of p_j, and u' is at distance $D \cdot 4^{-j}$ from ℓ (since we use an orthogonal vector to ℓ). Let $z' = f(u') - f(s')$. By definition of f, for all $i > j$ we have that $f(p_i \circ s) = f(s') + z'/4^{i-j}$, in particular $f(p_{k-1} \circ s) = f(s') + z'/4^{k-1-j}$, which suggests that the distance of $f(p_{k-1} \circ s)$ from the line ℓ is equal to $D \cdot 4^{-j}/4^{k-1-j} = D/4^{k-1}$.

As x is one of the four vertices created from the edge p_k whose end points are s' and $p_{k-1} \circ s$, it can be verified by the definition of the embedding f that its distance to ℓ is at least $1/4$ of the distance of $p_{k-1} \circ s$ to ℓ, that is at least $D/4^k$, as required.

Next we prove the inductive step. Let q, r be the level $m-1$ vertices which are the end points of the edge labeled p_m, and let ℓ' denote the line passing through $f(q)$ and $f(r)$. By the induction hypothesis the distance of both $f(q)$ and $f(r)$ from ℓ is at least $D \cdot (4^{-k} - \sum_{i=k+1}^{m-1} 4^{-i})$ and both are on the same side of ℓ. This suggests that every point on ℓ', in particular $p_m \circ s$ and $p_m \circ t$, is at least $D \cdot (4^{-k} - \sum_{i=k+1}^{m-1} 4^{-i})$ away from ℓ. It remains to argue about $p_m \circ u$ and $p_m \circ v$, which by definition are embedded by f at distance $D/4^m$ from ℓ', which means their distance to ℓ can be closer than that of $f(q), f(r)$ by at most $D/4^m$, which concludes the proof of (4).

Since $\sum_{i=k+1}^{m} 4^{-i} \le 4^k/2$ we have that $f(x)$ is at least $D \cdot 4^{-k}/2$ away from ℓ. If $j < k(y) < m$ is the minimal such that $y_{k(y)} \neq 0$ (and $k(y) = m$ if there is no such value), then a analogous argument will show that $f(y)$ is at least $D \cdot 4^{-k(y)}/2$ away from ℓ on the side of $f(v')$ where $v' = p_j \circ v$. In particular, this suggests that $\|f(x) - f(y)\| \ge D \cdot 4^{-k}/2$, as required.

□

It remains to bound the contraction for an arbitrary pair x, y.

Lemma 4. *For any $x, y \in V(G_n)$,*

$$\|(f \oplus g)(x) - (f \oplus g)(y)\| \ge D \cdot d(x,y)/128 \,. \tag{5}$$

Proof. First consider the case that $|g(x) - g(y)| \ge D \cdot d(x,y)/128$, then clearly (5) holds. Otherwise, $|g(x) - g(y)| < D \cdot d(x,y)/128$, and w.l.o.g assume that $g(x) < g(y)$. In this case, let $y' \in G_n$ be any point on a shortest path connecting y to a such that $g(x) = g(y')$. Then

$$d(y, y') = g(y) - g(y') = g(y) - g(x) \le D \cdot d(x,y)/128, \tag{6}$$

thus also

$$d(x, y') \ge d(x,y) - d(y,y') \ge 3d(x,y)/4 \,. \tag{7}$$

Using Lemma 3 on x, y' it follows that,

$$\begin{aligned} \|f(x) - f(y)\| &\ge \|f(x) - f(y')\| - \|f(y') - f(y)\| \\ &\overset{(3)}{\ge} D \cdot d(x,y')/32 - 2d(y',y) \\ &\overset{(6)\wedge(7)}{\ge} 3D \cdot d(x,y)/128 - D \cdot d(x,y)/64 \\ &= D \cdot d(x,y)/128 \,. \end{aligned}$$

□

The proof of Theorem 1 follows from Lemma 2 and Lemma 4

4 Embedding Doubling Metrics to Low Dimensional ℓ_∞

In this section we prove Theorem 2. Let us first remark that the dependence of the dimension D on the parameters is essentially tight (up to a constant in the exponent), that is $D \ge \lambda + (1/\epsilon)^{\Omega(1)} + \Omega(\log n)$: First, the $\log n$ term cannot be improved, because [GKL03] showed an $\Omega(\sqrt{\log n})$ lower bound on the distortion when embedding doubling metrics into ℓ_2. Under the ℓ_2 norm our embedding has distortion at most $(1+\epsilon)\sqrt{D}$, so when ϵ and λ are constants it must be that $D = \Omega(\log n)$. Second, a linear dependence on λ in the dimension is necessary, because for $\epsilon = 1$, say, the dimension of a normed space in which any n-point metric embeds with distortion 2 must be $\Omega(n) = \Omega(\lambda)$ [Mat02]. In the full version we show that there must be a polynomial dependence on $1/\epsilon$ as well.

4.1 Construction

For simplicity of presentation we first handle the case in which the spread (or aspect ratio) of the metric is at most n, that is, $\mathrm{diam}(X) < n$, the general case is deferred to Appendix A. For each $0 \le i < \log n$ take a r_i-net N_i, where $r_i = \epsilon \cdot 2^{i-2}$. Fix some net N_i, and for an integer $k > 0$ define a *spread-partition* $P_i(k)$ as a partition of N_i into k clusters $N_{i0}, N_{i1}, \dots, N_{i(k-1)}$, such that each cluster is well spread. Formally, for all $0 \le j \le k-1$, if $u, v \in N_{ij}$ then

$$d(u, v) \ge 5 \cdot 2^i \ . \tag{8}$$

Note that N_{ij} is not necessarily a net of N_i, as it may not satisfy the covering property of nets.

Claim 5. *Fix $k = \lambda^{6+\log(1/\epsilon)}$. For all $0 \le i < \log n$ there exists a spread-partition $P_i(k)$.*

Proof. To construct $P_i(k)$, first greedily choose a maximal $N_{i0} \subseteq N_i$ that satisfy (8). For any $0 < j \le k-1$, after choosing $N_{i0}, ..., N_{i(j-1)}$, greedily choose a maximal $N_{ij} \subseteq N_i \setminus (N_{i0} \cup ... \cup N_{i(j-1)})$ that satisfy (8). We claim that after k iterations N_i must be exhausted. Seeking contradiction, assume that $u \in N_i$ was not covered by any N_{ij}, and consider $B = B(u, 5 \cdot 2^i)$. By using the doubling property iteratively, the ball B can be covered by $\lambda^{\log(5 \cdot 2^i/(r_i/2))}$ balls of radius $r_i/2$, each of these small balls can contain at most one point from N_i. As $\lambda^{\log(5 \cdot 2^i/(r_i/2))} < k$, we conclude that for some $0 \le j \le k-1$, N_{ij} does not contain any point from B, but then by maximality it should have contained u, a contradiction. ☐

Next we define the embedding $f : X \to \mathbb{R}^D$ with $D = k \log n$, where k is defined as in Claim 5. Let $f_{ij}(x) = d(x, N_{ij})$, and

$$f(x) = \bigoplus_{i=1}^{\log n} \bigoplus_{j=0}^{k-1} f_{ij}(x) \ .$$

(We use the convention that if $N_{ij} = \emptyset$ then $d(x, N_{ij}) = 0$).

4.2 Proof

Fix some $x, y \in X$. By the triangle inequality we have that for any N_{ij}, $d(x, N_{ij}) - d(y, N_{ij}) \le d(x, y)$, so that for any $1 \le i \le \log n$ and $0 \le j \le k-1$, $f_{ij}(x) - f_{ij}(y) \le d(x, y)$. By symmetry of x, y this suggests that

$$|f_{ij}(x) - f_{ij}(y)| \le d(x, y) \ . \tag{9}$$

Next we show that there are i, j such that $f_{ij}(x) - f_{ij}(y) \ge d(x, y)(1 - \epsilon)$. Let $1 \le i \le \log n$ be such that $2^{i-1} \le d(x, y) < 2^i$, and let $0 \le j \le k-1$ be such that $d(x, N_{ij}) \le r_i$ (such a j must exist because N_i is an r_i-net). Denote by $u \in N_{ij}$ the point satisfying $d(x, N_{ij}) = d(x, u)$.

We claim that $d(y, N_{ij}) = d(y, u)$. To see this, first observe that $d(y, u) \le d(y, x) + d(x, u) \le 2^i + r_i < (5/4) \cdot 2^i$. Consider any other $v \in N_{ij}$, by the construction of N_{ij}, $d(v, u) \ge 5 \cdot 2^i$, so $d(y, v) \ge d(u, v) - d(y, u) > 5 \cdot 2^i - (5/4) \cdot 2^i > (5/4) \cdot 2^i > d(y, u)$. Thus it follows that $d(y, N_{ij}) = d(y, u) \ge d(y, x) - d(x, u) \ge d(x, y) - r_i$. We conclude that

$$f_{ij}(y) - f_{ij}(x) \ge (d(x, y) - r_i) - r_i = d(x, y) - \epsilon \cdot 2^{i-1} \ge d(x, y)(1 - \epsilon) \ . \tag{10}$$

The proof of Theorem 2 follows directly from (9) and (10).

5 Embedding Doubling Snowflakes

This section is devoted to the proof of Theorem 3. Recall that the snowflake of a metric (X,d) is the metric $(X,d^{1-\alpha})$ where $0<\alpha<1$. In the extremes, taking $\alpha=0$ gives the original metric and when $\alpha=1$ this is a uniform metric (all distances are 1). Observe that when $\alpha=\epsilon/(\log n)$ we have that $d^{1-\alpha}\approx d$ (up to a factor of $(1-\epsilon)$), which suggest that Theorem 3 is in fact an extension of Theorem 2. First we briefly mention why the dependence on α in the dimension D is tight. For $0<\alpha\le 1/2$ this can be seen by a result of [LMN05], who showed that the $1-\alpha$ snowflake of the Laakso graph (for which $\lambda=6$), must suffer $\Omega(\sqrt{1/\alpha})$ distortion when embedded into Euclidean space. Now, when ϵ is a constant our embedding has distortion $O(\sqrt{D})$ under the ℓ_2 norm, so it must be that $D=\Omega(1/\alpha)$ (the term $1/(1-\alpha)$ is just a constant in this case). In the other case $1/2<\alpha\le 1$, as α approaches 1 the metric becomes more uniform. In particular, when $\alpha=1-\gamma/\log n$ all distances raised to power α are between 1 and 2^γ. A simple volume argument suggests that metrics on n points with aspect ratio 2^γ, require $\Omega((\log n)/\gamma)=\Omega(1/(1-\alpha))$ dimensions for constant distortion embedding into ℓ_∞ (the term $1/\alpha$ is just a constant in this case).

5.1 Construction

Here too we shall assume first that $\operatorname{diam}(X)\le n$ (the general case is similar to the construction given in Appendix A and is deferred to the full version). We will also assume w.l.o.g that $1/(4\alpha)$ and $\alpha\log n$ are integers. For each $0\le i<\log n$, construct the nets N_i and spread partitions $P_i(k)$ exactly as in the previous section (recall that $r_i=\epsilon\cdot 2^{i-2}$ and $k=\lambda^{6+\log(1/\epsilon)}$). Next we define the embedding, fix $D=2k\log(1/\epsilon)/(\alpha(1-\alpha))$ where k is defined as in Claim 5 (assume that D is integer), and let $\{e_0,\dots,e_{D-1}\}$ be the standard orthonormal basis for $\mathbb{R}^D$, extended to an infinite sequence $\{e_j\}_{j\in\mathbb{N}}$ (that is, $e_j=e_{j(\text{mod }D)}$ for all $j\in\mathbb{N}$). For any $0\le i<\log n$ and $0\le j\le k-1$ let

$$g_{ij}(x)=\frac{\min\{2^i,d(x,N_{ij})\}}{2^{\alpha i}}$$

(we use the convention that if $N_{ij}=\emptyset$ then $d(x,N_{ij})=0$). Define the embedding $f:X\to\mathbb{R}^D$ by

$$f(x)=\sum_{i=0}^{\log n}\sum_{j=0}^{k-1}g_{ij}(x)\cdot e_{ik+j}\ .$$

Consider the h-th coordinate of the embedding f_h, with $0\le h\le D-1$. Observe that there is a unique value of $0\le j\le k-1$ such that $ik+j=h(\text{mod }D)$ could hold for some i, and let $j(h)$ be that value. Letting $I(h)=\{i\ :\ ik+j(h)=h(\text{mod }D)\}$ we have that

$$f_h(x)-f_h(y)=\sum_{i\in I(h)}g_{ij}(x)-g_{ij}(y)\ . \tag{11}$$

We may enumerate $I(h)=\dots,i_{-1},i_0,i_1,\dots$ such that $i_0\in I(h)$ is some fixed scale, and $i_s=i_0+2s\log(1/\epsilon)/(\alpha(1-\alpha))$ for all $s\in\mathbb{Z}$. In what follows we show that f has distortion $1+O(\epsilon)$.

5.2 Expansion Bound

Here we show that the embedding f under the ℓ_∞ norm does not expand distances by more than a factor of $1+3\epsilon$. Fix a pair $x,y\in X$, and observe that by the triangle inequality we have that

$$g_{ij}(x)-g_{ij}(y)\le d(x,y)\cdot 2^{-i\alpha}\ , \tag{12}$$

and also

$$g_{ij}(x)-g_{ij}(y)\le 2^{i(1-\alpha)}\ , \tag{13}$$

for all $0 \le i < \log n$ and $0 \le j \le k-1$. Consider the h-th coordinate of the embedding f_h, with $0 \le h \le D-1$, and fix $j = j(h)$. Let $i_0 \in I(h)$ be the maximal scale in $I(h)$ such that $2^{i_0} \le d(x,y)$. First we bound the small distance scales appearing in $I(h)$,

$$\begin{aligned}
\sum_{i \in I(h)\,:\, i < i_0} g_{ij}(x) - g_{ij}(y) &\le \sum_{s=-\infty}^{-1} g_{i_s j}(x) - g_{i_s j}(y) \\
&\overset{(13)}{\le} \sum_{s=1}^{\infty} 2^{i_{-s}(1-\alpha)} \\
&= \sum_{s=1}^{\infty} 2^{(1-\alpha)(i_0 - 2s\log(1/\epsilon)/(\alpha(1-\alpha)))} \\
&\le 2^{i_0(1-\alpha)} \sum_{s=1}^{\infty} \epsilon^{2s/\alpha} \\
&= 2^{i_0(1-\alpha)} \frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}} \,. \qquad (14)
\end{aligned}$$

Next we bound the contribution from the high distance scales,

$$\begin{aligned}
\sum_{i \in I(h)\,:\, i > i_0} g_{ij}(x) - g_{ij}(y) &\le \sum_{s=1}^{\infty} g_{i_s j}(x) - g_{i_s j}(y) \\
&\overset{(12)}{\le} \sum_{s=1}^{\infty} d(x,y) \cdot 2^{-i_s \alpha} \\
&= d(x,y) \sum_{s=1}^{\infty} 2^{-i_0\alpha - 2s\log(1/\epsilon)/(1-\alpha)} \\
&= d(x,y) \cdot 2^{-i_0\alpha} \sum_{s=1}^{\infty} \epsilon^{2s/(1-\alpha)} \\
&\le d(x,y) \cdot 2^{-i_0\alpha} \frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}} \,. \qquad (15)
\end{aligned}$$

For the critical scale i_0 we have by (13) that

$$g_{i_0 j}(x) - g_{i_0 j}(y) \le 2^{i_0(1-\alpha)} \,. \qquad (16)$$

Combining the bounds of (14), (15) and (16), we get

$$f_h(x) - f_h(y) \le 2^{i_0(1-\alpha)} \left(1 + \frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}}\right) + d(x,y) \cdot 2^{-i_0\alpha} \frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}} \qquad (17)$$

It remains to show that the RHS of (17) is bounded by $d(x,y)^{1-\alpha}(1+3\epsilon)$. We verify this using a simple case analysis, and write the calculations for completeness. Note that by maximality of i_0 we have that $\epsilon^{2/(\alpha(1-\alpha))} \cdot d(x,y) \le 2^{i_0} \le d(x,y)$. If it is the case that $d(x,y) \cdot \epsilon^{1/(\alpha(1-\alpha))} \le 2^{i_0} \le d(x,y)$, then intuitively, the dominant term is the one coming from (16), whereas the terms in (14) and (15) will contributes only an ϵ-fraction of that. Formally, we bound the RHS of (17) by

$$\begin{aligned}
&2^{i_0(1-\alpha)}\left(1+\frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}}\right)+d(x,y)\cdot 2^{-i_0\alpha}\frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\\
\leq\ & d(x,y)^{1-\alpha}\left(1+\frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}}\right)+\frac{d(x,y)^{1-\alpha}}{\epsilon^{1/(1-\alpha)}}\cdot\frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\\
\leq\ & d(x,y)^{1-\alpha}\left(1+\epsilon\right)+d(x,y)^{1-\alpha}\cdot\frac{\epsilon^{1/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\\
\leq\ & d(x,y)^{1-\alpha}(1+3\epsilon)\ .
\end{aligned}$$

The last inequalities are using that $\epsilon<1/2$ so that $\frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\leq\epsilon^2(1+\epsilon^2)\leq\epsilon$ and $\frac{\epsilon^{1/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\leq\epsilon(1+\epsilon)\leq 2\epsilon$. The second case where $d(x,y)\cdot\epsilon^{2/(\alpha(1-\alpha))}\leq 2^{i_0}<d(x,y)\cdot\epsilon^{1/(\alpha(1-\alpha))}$ is similar, here the dominant term will come from (15):

$$\begin{aligned}
&2^{i_0(1-\alpha)}\left(1+\frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}}\right)+d(x,y)\cdot 2^{-i_0\alpha}\frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\\
\leq\ & d(x,y)^{1-\alpha}\cdot\epsilon^{1/\alpha}\cdot\left(1+\frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}}\right)+\frac{d(x,y)^{1-\alpha}}{\epsilon^{2/(1-\alpha)}}\cdot\frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\\
\leq\ & d(x,y)^{1-\alpha}\cdot\epsilon(1+\epsilon)+d(x,y)^{1-\alpha}\cdot\epsilon\\
\leq\ & d(x,y)^{1-\alpha}(1+3\epsilon)\ .
\end{aligned}$$

We have shown that $f_h(x)-f_h(y)$ is bounded, by symmetry of x,y also $|f_h(x)-f_h(y)|\leq d(x,y)^{1-\alpha}\cdot(1+3\epsilon)$ as well, which concludes the expansion bound.

5.3 Contraction Bound

Finally we bound the contraction of the embedding. Fix a pair $x,y\in X$. We will show that there exist a single coordinate $0\leq h\leq D-1$ such that $|f_h(x)-f_h(y)|\geq d(x,y)^{1-\alpha}\cdot(1-5\epsilon)$. Let $0\leq i_0\leq\log n$ be such that $2^{i_0-1}<d(x,y)\leq 2^{i_0}$, and let $0\leq j\leq k-1$ be such that $d(x,N_{i_0j})\leq r_{i_0}$. Let $u\in N_{i_0j}$ be the point satisfying $d(x,u)=d(x,N_{i_0j})$, and the same calculation as in Section 4.2 shows that $d(y,N_{i_0j})\geq d(x,y)-r_i$, thus

$$g_{i_0j}(y)-g_{i_0j}(x)\geq d(x,y)(1-\epsilon)\ . \tag{18}$$

Let $0\leq h\leq D-1$ be such that $h=i_0k+j(\bmod D)$, for the values of i_0,j fixed above. The scale i_0 is the critical scale which by (18) provides sufficient contribution for x,y, and it remains to show that the other scales participating in coordinate h do not cancel out this contribution. By (14) we have that

$$\begin{aligned}
\sum_{i\in I(h)\,:\,i<i_0} g_{ij}(x)-g_{ij}(y) &\leq 2^{i_0(1-\alpha)}\frac{\epsilon^{2/\alpha}}{1-\epsilon^{2/\alpha}}\\
&\leq (2d(x,y))^{1-\alpha}\epsilon(1+\epsilon)\\
&\leq d(x,y)^{1-\alpha}\cdot 3\epsilon\ .
\end{aligned}\tag{19}$$

By (15) we have that

$$\begin{aligned}
\sum_{i\in I(h)\,:\,i>i_0} g_{ij}(x)-g_{ij}(y) &\leq d(x,y)\cdot 2^{-i_0\alpha}\frac{\epsilon^{2/(1-\alpha)}}{1-\epsilon^{2/(1-\alpha)}}\\
&\leq d(x,y)^{1-\alpha}\cdot\epsilon\ .
\end{aligned}\tag{20}$$

Combining (19) and (20) with the contribution of the critical scales i_0 in (18) we obtain that

$$\begin{aligned} |f_h(x) - f_h(y)| &\geq |g_{i_0 j}(x) - g_{i_0 j}(y)| - \left| \sum_{i \in I(h)\,:\, i \neq i_0} g_{ij}(x) - g_{ij}(y) \right| \\ &\geq (1 - 5\epsilon) d(x,y)^{1-\alpha} . \end{aligned}$$

6 acknowledgements

The author is grateful to Moses Charikar, Michael Elkin and Lee-Ad Gottlieb for fruitful discussions. This work is supported by ISF grant No. (523/12) and by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n°303809.

References

[ABN08] Ittai Abraham, Yair Bartal, and Ofer Neiman. Embedding metric spaces in their intrinsic dimension. In *Proceedings of the 19th annual ACM-SIAM symposium on Discrete algorithms*, SODA '08, pages 363–372, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[Ass83] P. Assouad. Plongements lipschitziens dans $\mathbb{R}^n$. *Bull. Soc. Math. France*, 111(4):429–448, 1983.

[BC05] Bo Brinkman and Moses Charikar. On the impossibility of dimension reduction in l_1. *J. ACM*, 52(5):766–788, 2005.

[BGK12] Yair Bartal, Lee-Ad Gottlieb, and Robert Krauthgamer. The traveling salesman problem: low-dimensionality implies a polynomial time approximation scheme. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 663–672, New York, NY, USA, 2012. ACM.

[BRS11] Yair Bartal, Ben Recht, and Leonard J. Schulman. Dimensionality reduction: beyond the johnson-lindenstrauss bound. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 868–887. SIAM, 2011.

[CGMZ05] T-H. Hubert Chan, Anupam Gupta, Bruce M. Maggs, and Shuheng Zhou. On hierarchical routing in doubling metrics. In *Proc. 16th Ann. ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2005.

[CGT10] T.-H. Hubert Chan, Anupam Gupta, and Kunal Talwar. Ultra-low-dimensional embeddings for doubling metrics. *J. ACM*, 57(4), 2010.

[CK10] Jeff Cheeger and Bruce Kleiner. Differentiating maps into L^1, and the geometry of BV functions. *Annals of Math.*, 171(2):1347–1385, 2010.

[CKN09] Jeff Cheeger, Bruce Kleiner, and Assaf Naor. A $(\log n)^{\Omega(1)}$ integrality gap for the sparsest cut sdp. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '09, pages 555–564, Washington, DC, USA, 2009. IEEE Computer Society.

[GK11] Lee-Ad Gottlieb and Robert Krauthgamer. A nonlinear approach to dimension reduction. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, pages 888–899. SIAM, 2011.

[GKL03] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '03, pages 534–543, Washington, DC, USA, 2003. IEEE Computer Society.

[GR08] Lee-Ad Gottlieb and Liam Roditty. Improved algorithms for fully dynamic geometric spanners and geometric routing. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '08, pages 591–600, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[HPM06] Sariel Har-Peled and Manor Mendel. Fast construction of nets in low-dimensional metrics and their applications. *SIAM Journal on Computing*, 35(5):1148–1184, 2006.

[JLM09] Alex Jaffe, James R. Lee, and Mohammad Moharrami. On the optimality of gluing over scales. In *Proceedings of the 12th International Workshop and 13th International Workshop on Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, APPROX '09 / RANDOM '09, pages 190–201, Berlin, Heidelberg, 2009. Springer-Verlag.

[KLMN05] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: a new embedding method for finite metrics. *Geometric and Functional Analysis*, 15(4):839–858, 2005.

[Laa02] Tomi J. Laakso. Plane with A_∞-weighted metric not bi-Lipschitz embeddable to $\mathbb{R}^N$. *Bull. London Math. Soc.*, 34(6):667–676, 2002.

[LLR95] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[LMN05] James R. Lee, Manor Mendel, and Assaf Naor. Metric structures in l1: dimension, snowflakes, and average distortion. *Eur. J. Comb.*, 26(8):1180–1190, November 2005.

[LN04] James R. Lee and Assaf Naor. Embedding the diamond graph in l_p and dimension reduction in l_1. pages 745–747, 2004.

[LP01] Urs Lang and Conrad Plaut. Bilipschitz embeddings of metric spaces into space form. *Geometriae Dedicata*, 87(1-3):285–307, 2001.

[LS11] James R. Lee and Anastasios Sidiropoulos. Near-optimal distortion bounds for embedding doubling spaces into l1. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 765–772, New York, NY, USA, 2011. ACM.

[Mat02] Jiri Matoušek. *Lectures on discrete geometry*. Springer-Verlag, New York, 2002.

[NN12] Assaf Naor and Ofer Neiman. Assouad's theorem with dimension independent of the snowflaking. *Revista Matematica Iberoamericana*, 28(4,):1–21, 2012.

[Sem96] S. Semmes. On the nonexistence of bilipschitz parameterizations and geometric problems about a_∞ weights. *Revista Matemática Iberoamericana*, 12:337–410, 1996.

A Proof of Theorem 2 for Metrics with Arbitrary Diameter

Here we complete the proof of Theorem 2, without the restriction on the diameter of the metric. The technique is quite standard and appeared before. The idea is to allocate only $O(\log n)$ coordinates, and re-use the same coordinates every $O(\log n)$ scales.

A.1 Construction

Let $\Delta = \mathrm{diam}(X)$ and for all $0 \le i \le \log\Delta$ let (X, d_i) be the metric induced by contracting distances smaller than $2^i/n^2$. Formally, consider the complete graph on vertices X with edge $\{u, v\}$ having weight $d(u, v)$. Replace all weights smaller than $2^i/n^2$ by 0, and let d_i be the shortest path metric on this graph. It can be checked that $d(x, y) - 2^i/n \le d_i(x, y) \le d(x, y)$ for all $x, y \in X$.

For each $0 \le i \le \log\Delta$ take a r_i-net N_i with respect to (X, d_i), where $r_i = \epsilon \cdot 2^{i-2}$. Create a spread-partition $P_i(k) = \{N_{i0}, N_{i1}, \dots, N_{i(k-1)}\}$ (recall that $k = \lambda^{6+\log(1/\epsilon)}$). Next we define the embedding, fix $D = 4k \log n$, and let $\{e_0, \ldots, e_{D-1}\}$ be the standard orthonormal basis for $\mathbb{R}^D$, extended to an infinite sequence $\{e_j\}_{j \in \mathbb{N}}$ (that is, $e_j = e_{j \pmod D}$ for all $j \in \mathbb{N}$). For any $0 \le i \le \log \Delta$ and $0 \le j \le k-1$ let

$$g_{ij}(x) = \min\{2^{i+1}, d_i(x, N_{ij})\} .$$

Define the embedding $f : X \to \mathbb{R}^D$ by

$$f(x) = \sum_{i=0}^{\log \Delta} \sum_{j=0}^{k-1} g_{ij}(x) \cdot e_{ik+j} .$$

A.2 Expansion Bound

Now we show that the embedding f under the ℓ_∞ norm does not expand distances by more than a factor of $1 + 1/n$. Fix a pair $x, y \in X$, and consider the h-th coordinate of the embedding f_h, with $0 \le h \le D-1$. We have that $f_h(x) - f_h(y) = \sum_{i,j \,:\, h = ik+j (\mathrm{mod}\ D)} g_{ij}(x) - g_{ij}(y)$. Let $0 \le i' \le \log \Delta$ be such that $2^{i'-1} \le d(x,y) < 2^{i'}$, then for all $i > i' + 2 \log n$ it holds that $d(x,y) < 2^i/n^2$ and thus $d_i(x,y) = 0$, in particular, $g_{ij}(x) = g_{ij}(y)$ and so there is no contribution at all from such scales. By the triangle inequality we also have that $g_{ij}(x) - g_{ij}(y) \le d_i(x,y)$ and also $g_{ij}(x) - g_{ij}(y) \le 2^{i+1}$ for all $0 \le i \le \log \Delta$.

$$\begin{aligned} f_h(x) - f_h(y) &\le \sum_{i,j \,:\, i \le i' + 2\log n, h = ik+j (\mathrm{mod}\ D)} g_{ij}(x) - g_{ij}(y) \\ &\le \sum_{i,j \,:\, i' - 2\log n < i \le i' + 2\log n, h = ik+j (\mathrm{mod}\ D)} g_{ij}(x) - g_{ij}(y) \\ &\quad + \sum_{i \le i' - 2\log n} 2^{i+1} \\ &\le d_i(x,y) + 2^{i'+2}/n^2 \\ &\le d(x,y)(1 + 1/n) . \end{aligned}$$

The third inequality holds as there is at most one value of i with $i' - 2\log n < i < i' + 2 \log n$ such that $h = ik + j$.

A.3 Contraction Bound

Finally we bound the contraction of the embedding. Fix a pair $x, y \in X$. We will show that there exist a single coordinate $0 \le h \le D-1$ such that $|f_h(x) - f_h(y)| \ge (1-\epsilon) d(x,y)$. Let $0 \le i \le \log \Delta$ such that $2^i \le d(x,y) < 2^{i+1}$, and let $0 \le j \le k-1$ be such that $d_i(x, N_{ij}) \le r_i$ (such a j must exist because N_i is an r_i-net). Denote by $u \in N_{ij}$ the point satisfying $d_i(x, N_{ij}) = d_i(x,u)$, then since $\epsilon < 1$ also $g_{ij}(x) \le r_i$.

We claim that $d_i(y, N_{ij}) = d_i(y,u)$. To see this, first observe that $d_i(y,u) \le d_i(y,x) + d_i(x,u) \le 2^{i+1} + r_i < (5/4) \cdot 2^{i+1}$. Consider any other $v \in N_{ij}$, by the construction of N_{ij}, $d_i(v,u) \ge 5 \cdot 2^i$, so $d_i(y,v) \ge d_i(u,v) - d_i(y,u) > (5/2) \cdot 2^{i+1} - (5/4) \cdot 2^{i+1} = (5/4) \cdot 2^{i+1} > d_i(y,u)$. Thus it follows that $g_{ij}(y) \ge d_i(y,u) \ge d_i(y,x) - d_i(x,u) \ge d_i(x,y) - r_i$. We conclude that

$$g_{ij}(y) - g_{ij}(x) \ge (d_i(x,y) - r_i) - r_i = d_i(x,y) - \epsilon \cdot 2^{i-1} \ge d(x,y)(1 - \epsilon/2 - 1/n) .$$

Let $0 \le h \le D-1$ be such that $h = ik + j (\mathrm{mod} D)$, for the values of i, j fixed above. Then we claim that any other pair i', j such that $h = i'k + j (\mathrm{mod} D)$ has either 0 or very small contribution to the h coordinate. If $i' > i$ then it must be that $i' \ge 4 \log n \cdot i$ so that $d(x,y) \le 2^{i+1} < 2^{i'}/n^2$, thus as before $g_{i'j}(x) = g_{i'j}(y)$. For values of i' such that $i' < i$ then $i' \leq i - 4\log n$ thus

$$\begin{aligned}\sum_{i'<i,j\,:\,h=i'k+j(\text{mod }D)} |g_{i'j}(x) - g_{i'j}(y)| &\leq \sum_{i'\leq i-4\log n} 2^{i'+1} \\ &\leq 2^i/n^2 \\ &\leq d(x,y)/n^2 \ .\end{aligned}$$

Finally,

$$\begin{aligned}\|f(x) - f(y)\|_\infty &\geq |f_h(x) - f_h(y)| \\ &\geq |g_{ij}(y) - g_{ij}(x)| - \sum_{i'<i,j\,:\,h=i'k+j(\text{mod }D)} |g_{i'j}(x) - g_{i'j}(y)| \\ &\geq d(x,y)(1 - \epsilon/2 - 2/n) \ .\end{aligned}$$

If $10/n \leq \epsilon \leq 1/4$ then the distortion is indeed $(1 + 1/n)/(1 - \epsilon/2 - 2/n) \leq 1 + \epsilon$.