Bounded geometries, fractals, and low-distortion embeddings

Anupam Gupta[*]   Robert Krauthgamer[†]   James R. Lee[‡]

Abstract

The doubling constant *of a metric space* (X, d) *is the smallest value* λ *such that every ball in* X *can be covered by* λ *balls of half the radius. The* doubling dimension *of* X *is then defined as* $\dim(X) = \log_2 \lambda$. *A metric (or sequence of metrics) is called* doubling *precisely when its doubling dimension is bounded. This is a robust class of metric spaces which contains many families of metrics that occur in applied settings.*

We give tight bounds for embedding doubling metrics into (low-dimensional) normed spaces. We consider both general doubling metrics, as well as more restricted families such as those arising from trees, from graphs excluding a fixed minor, and from snowflaked metrics. Our techniques include decomposition theorems for doubling metrics, and an analysis of a fractal in the plane due to Laakso [21]. Finally, we discuss some applications and point out a central open question regarding dimensionality reduction in L_2.

1 Introduction

A basic goal in the study of finite metric spaces is to approximate some class of metric spaces by another more simple or tractable class. Apart from being beautiful objects of study lying at the intersection of analysis, combinatorics, and geometry, the ideas and techniques generated in this field have led to a number of powerful algorithmic applications (see e.g. [15, 25, 29]).

We consider embeddings of finite metric spaces into L_p spaces. Given a metric (X, d), the goal is to find a map $f : X \to L_p$ such that $||f(x) - f(y)||_p$ is close to $d(x, y)$ for all $x, y \in X$. The worst-case factor by which distances are expanded or contracted is called the *distortion* of the map f. In general, our goal is to find bounds on the achievable

distortion in terms of certain fundamental properties of the metric (X, d).

The general case is well-understood. Bourgain [4] showed that every n-point metric embeds into L_p with $O(\log n)$ distortion for any fixed p; it is shown in [26] that this bound is tight, for all $p \le 2$, for the shortest path metric on constant-degree expander graphs. This was later extended in [27], showing a tight upper bound of $O\left(\frac{\log n}{p}\right)$ for any L_p space.

In light of this, a significant amount of effort has been made to understand the distortion achievable for restricted classes of metric spaces. So far, the restrictions considered have been mostly topological. For L_1 embeddings, this is due partly to the intimate connection with multicommodity flows and approximations for the sparsest cut, see e.g. [26, 3, 13, 33]. It is not too difficult to see that every tree metric embeds isometrically into L_1. Matoušek [28] showed that every tree embeds into L_p with distortion $O((\log \log n)^{\min(\frac{1}{2}, \frac{1}{p})})$ and that this bound is tight for all $p > 1$. Rao [34] showed that every planar graph embeds into L_2 with distortion $O(\sqrt{\log n})$, and this in fact holds for any family which excludes a fixed minor. A matching lower bound, yielded by a family of series-parallel metrics was given in [30] (see also [20, 21, 22]). Gupta et al. [13] show that K_4-free (series-parallel) and $K_{2,3}$-free (outerplanar) graphs embed into L_1 with constant distortion.

Here, we consider restrictions not on the topology of the metric, but on its geometry. More specifically, we examine how the "volume growth" of a metric affects its embeddability into L_p spaces. The notion of growth that we use is well-studied, and is very similar to a notion of Assouad [2], see also [14]. Our definition is technically slightly different from Assouad's, but the flavor is left unaltered; in particular, the notion of bounded growth is equivalent under either framework.

For a metric (X, d), let the *doubling constant* be be the smallest value λ such that every ball in X can be covered by λ balls of half the radius, and define the *doubling dimension* of X as $\dim(X) = \log_2 \lambda$. It can be shown, for instance, that for every fixed p, the doubling dimension of d-dimensional ℓ_p space is roughly d, and clearly for a finite metric $\dim(X) \le \log |X|$. A metric (or sequence of metrics) is called *doubling* precisely when its doubling dimen-

sion is bounded. We discuss this notion more thoroughly in Section 1.3.

Not only are doubling metrics interesting objects in their own right, but they are also of practical concern. Growth restrictions are very natural and are thought to occur in real-world phenomena such as peer-to-peer networks (see e.g. [31]) and data analysis (e.g., when the input data resides on a low-dimensional manifold, cf. [36]). In fact, various algorithms can be tailored to run efficiently on certain classes of growth-restricted metrics, as demonstrated in [9, 32, 17, 19]. The metrics considered there are either equivalent to or a subclass of those metrics which are doubling; see Section 1.3.

1.1 Results and techniques

We are concerned with the broad roles of "volume" and "structure" in determining the embeddability of a metric. For instance, in Section 2, we show that every doubling tree metric admits a constant distortion embedding into $\ell_p^{O(1)}$ (i.e., constant dimensional ℓ_p space) for any $p \in [1, \infty]$. This exhibits a very natural class of metric spaces which embed into ℓ_2 with $O(1)$ distortion, but not isometrically. As discussed before, some tree metrics require $\Omega(\sqrt{\log \log n})$ distortion to embed into ℓ_2 [5, 28], while we prove that some doubling metrics require distortion $\Omega(\sqrt{\log n})$ (see Section 5). Thus it is precisely the synthesis of these two properties that yields an enormous improvement in embeddability.

That these metrics can be embedded using only $O(1)$ dimensions is perhaps even more surprising (see the discussion in Section 2), and as we will soon see, proves a special case of a conjecture of Assouad. Our embedding makes use of a novel partitioning algorithm for doubling trees. The partition is described conveniently by a coloring of the edges. Our algorithm either finds a good coloring or a submetric which is a counterexample to the doubling assumption. To achieve an embedding into $O(1)$ dimensions, we must reuse colors. We do this by defining a notion of adjacency between paths, and arguing that the resulting graph (whose vertices are paths) has bounded chromatic number.

In Section 3, we construct low-diameter decompositions for general metrics whose parameters depend smoothly on the doubling dimension. Such decompositions are the main tool in many embedding results, as well as a number of other applications. To construct these decompositions, we adapt a probabilistic technique of [7]. For applications of the decomposition later in the paper, it is important that the probability space that we sample from be very compactly defined (e.g., of size $O(1)$ for the case of doubling metrics). For this purpose, we use some ideas from [18], in conjunction with Lovász Local Lemma, in order to exploit certain locality properties of our decomposition. Our use of the local lemma here, and elsewhere in the paper, can be made

algorithmic using standard techniques.

In Section 4, we devise a number of embeddings, some of which rely heavily on the clustering of Section 3. First, we give a Bourgain-style embedding which shows that every general doubling metric (X, d) embeds into ℓ_p with distortion $O((\log n)^{\min(\frac{1}{2}, \frac{1}{p})})$ for any $p \in [1, \infty)$. Instead of forming coordinates by taking the distance to arbitrary subsets of X, we instead use only subsets of appropriately sized *nets* in X (see Section 1.2 for the definition of a net). Using this embedding, along with another application of the local lemma, we show that every doubling metric can be embedded with $1 + \epsilon$ distortion into $\ell_\infty^{O(\log n)}$. See Section 6 for an application to distance labeling schemes.

Although this Bourgain-style embedding is fairly simple, its distortion degrades badly (exponentially) with $\dim(X)$. Based on the clustering of Section 3, and an embedding technique of Rao [34], we show that every n-point metric (X, d) embeds into ℓ_p with distortion at most

$$O\left(\dim(X) \cdot (\log n)^{\min(\frac{1}{2}, \frac{1}{p})}\right).$$

This provides a smooth upper bound on the distortion of general metrics in terms of their dimension. In the case of Euclidean embeddings, for instance, it beats Bourgain's general upper bound as long as $\dim(X) = o(\sqrt{\log n})$.

In [2], it is proved that if (X, d) is a doubling metric, then for any $0 < \alpha < 1$, (X, d^α) embeds into ℓ_2^k with distortion D, where k and D depend only on the doubling dimension of X. Here, (X, d^α) is the metric with all distances raised to the power α (this metric is called a *snowflaked* version of X). Unfortunately, the dependence of k and D on $\dim(X)$ is exponential. In Section 4, using a slight twist on the above embedding technique, we give an algorithmic version of Assouad's proof, and drastically improve the dependence of k and D to near-linear. Assouad also conjectured that the above result holds even when $\alpha = 1$. Although Semmes [35] disproved this conjecture, we have shown that it holds whenever (X, d) is a doubling tree metric.

In Section 5, we exhibit a family of series-parallel doubling metrics which requires $\Omega(\sqrt{\log n})$ distortion to embed into ℓ_2, which shows that the upper bound of Section 4 is tight. This family is based on the construction of a fractal due to Laakso [21]. (It was brought to our attention that a similar analysis was obtained independently in [22], but without an explicit dependence on n. In addition, the proof techniques are subtly different.)

Finally, in Section 6, we mention some algorithmic applications of our results, and point out a very intriguing open question regarding the roles of volume and structure in dimensionality reduction in Euclidean spaces.

1.2 Preliminaries

Here are some definitions used in the paper; the books by Deza and Laurent [10] and by Heinonen [14] give more details on metric spaces. [1] Let (X, d_X) and (Y, d_Y) be two metric spaces, and consider an injective map $f : X \to Y$. We define

$$\text{contraction}(f) = \sup_{a,b \in X} \frac{d_X(a,b)}{d_Y(f(a), f(b))},$$

$$\text{expansion}(f) = \sup_{a,b \in X} \frac{d_Y(f(a), f(b))}{d_X(a,b)}.$$

The *distortion* of f is defined by $\text{dist}(f) = \text{expansion}(f) \cdot \text{contraction}(f) = \|f\|_{\text{Lip}} \cdot \|f^{-1}\|_{\text{Lip}}$. The distortion with which X embeds into Y is the infimum of $\text{dist}(f)$ over all injective maps $f : X \to Y$.

As usual, we define $c_p(X)$ as the least distortion with which X embeds into some ℓ_p space. Let $G = (V, E)$ be a simple undirected graph with non-negative edge lengths. The edge lengths on G induce a natural metric $d_G(u, v)$ given by the length of a shortest path between u and v in G.

If we have two maps $f_1 : X \to Y_1$ and $f_2 : X \to Y_2$, we define their *direct sum* $f_1 \oplus f_2$ as the map $f : X \to Y_1 \times Y_2$ given by $f(x) = (f_1(x), f_2(x))$. This extends naturally to a direct sum of more than two functions.

We define the *aspect ratio* of (X, d) to be the ratio of the largest distance to the smallest in X. For $x \in X$ and $r \geq 0$, we define the *open ball of radius r about x* to be $B(x, r) = \{y \in X : d(x, y) < r\}$.

Finally, we say that a subset Y of X is an *ϵ-net* if it satisfies (1) For every $x, y \in Y, d(x, y) \geq \epsilon$ and (2) $X \subseteq \bigcup_{y \in Y} B(y, \epsilon)$. Such nets always exist for any $\epsilon > 0$. For finite metrics, they can be constructed greedily. For arbitrary metrics, proof of their existence is an easy application of Zorn's lemma.

1.3 Notions of dimension and volume

Here we consider some notions meant to capture the "volume growth" of arbitrary metric spaces and discuss relations between them.

Doubling dimension. An alternative definition for the doubling constant considers diameter (of subsets) instead of radius (of balls); that is, the doubling constant of a metric space (X, d) is as the smallest λ such that every subset of X of diameter $2r$ can be covered by at most λ subsets of diameter at most r. Again, the doubling dimension of X is then defined as $\dim(X) = \log_2 \lambda$. The main advantage of this definition is that for any submetric $Y \subseteq X$, we have $\dim(Y) \leq \dim(X)$. In what follows, we will find it easier

to work with balls rather than arbitrary sets. Thus we will use the definition given earlier (i.e., every ball in X can be covered by λ balls of half the radius). It is easy to see that moving between the two definitions affects the dimension by only a factor of 2.

The seminal paper of Assouad [2] showed that this notion attains several natural properties (see also [14, Ch. 10]). For instance, it can be shown that $\dim(\mathbb{R}^k) = \Theta(k)$ when $\mathbb{R}^k$ is endowed with the ℓ_p norm. It follows that the upper bound $O(k)$ applies to any subset of $\mathbb{R}^k$. For a finite metric space, clearly $\dim(X) \leq \log |X|$. The next proposition will be key (throughout); it is proved by applying the definition several times to obtain a cover with balls of sufficiently small radius, and then arguing that each net point is covered by a distinct ball.

Proposition 1.1. *Let (X, d) be a metric with doubling constant λ. If all pairwise distances in $Y \subseteq X$ are at least r (e.g., Y is an r-net of X), then for any point $x \in X$ and radius $t \geq r$ we have $|B(x, t) \cap Y| \leq \lambda^{\lceil \log \frac{2t}{r} \rceil}$.*

The counting measure. Karger and Ruhl [17] considered a notion of dimension that relies on the counting measure (in finite metric spaces): Let K be the smallest constant such that $|B(x, 2r)| \leq K|B(x, r)|$ for all $x \in X$, $r \geq 0$. (Strictly speaking, the requirement in their definition was enforced only when $|B(x, r)|$ was at least some threshold n_0. We ignore this for simplicity, but most of our techniques are local and allow for such restrictions.) Throughout, we refer to K as the *KR-constant* of X and define the *KR-dimension* of X as $\dim_{\text{KR}}(X) = \log_2 K$. The next proposition shows that bounded KR-dimension is a more stringent requirement than bounded doubling dimension. Its proof is relatively simple and is omitted from this version. However, bounded KR-dimension means that the counting measure is doubling, and thus this proposition is no more than an instantiation of a well-known basic result: If a metric space (X, d) has a doubling measure then the metric is doubling. (See e.g. [14] for a thorough treatment of doubling measures.)

Proposition 1.2. *For any finite metric (X, d), $\dim(X) \leq 4 \cdot \dim_{\text{KR}}(X)$.*

The converse, however, is not true; there are metrics with bounded Assouad dimension whose KR-dimension is $\Omega(\log |X|)$. For example, take an integer grid in the Euclidean plane, and consider only the origin and the points in the annulus $n/2 < \|x\| \leq n$. This reflects certain frailties in the definition of KR-dimension: Even if $\dim_{\text{KR}}(X)$ is bounded, this does not necessarily hold for submetrics of X. Furthermore, the annulus itself has a bounded KR-dimension, but this property is not maintained when even one point (the origin) is added.

---

[1] We are concerned mostly with finite metrics, but most of our results extend to arbitrary metric spaces via standard compactness arguments.

Local density. Finally, there is another natural notion of volume, which has been used widely in the study of the bandwidth of graphs [8, 11]. Given an *unweighted* connected graph $G = (V, E)$, the local density of G, denoted $\beta(G)$, is the smallest value β such that $|B(v, r)| \leq \beta r$ for all $v \in V, r > 0$. It is easy to see that since G is unweighted and connected, $|B(v, r)| \geq r$, hence $|B(v, 2r)| \leq 2\beta r \leq 2\beta|B(v, r)|$, which implies that $\dim(G) \leq 4 \dim_{KR}(G) \leq 4(\log \beta(G) + 1)$.

2 Trees

In this section we prove Assouad's conjecture for trees, showing that every doubling tree can be embedded with *constant distortion* into *constant dimensional* ℓ_p space for every $p \in [1, \infty]$, where both constants depend only on the doubling constant of the tree. Not many families of metrics are known to have such powerful embeddings, and it may be surprising that two seemingly orthogonal constraints are needed for such a result. However, as mentioned in the introduction, a constant distortion Euclidean embedding cannot be achieved if either of the constraints is dropped. Furthermore, the tree requirement cannot be relaxed either topologically (i.e., to graphs excluding a fixed minor) or graph-theoretically (i.e., to bounded treewidth graphs), as even series-parallel graphs might require $\Omega(\sqrt{\log n})$ distortion in any Euclidean embedding (see Section 5).

The fact that these results can be extended to constant dimension is even more surprising, especially since recent results of Brinkman and Charikar [6] (see also a short proof of [24], which even generalizes to the metrics exhibited in Section 5 [23]) show extremely strong lower bounds on the dimension required to embed simple series-parallel graphs into ℓ_1.

On a high level, our embeddings consist of two steps. Without loss of generality, we can assume that the tree is rooted at some vertex. The first step partitions the edges of the tree into monotone paths (i.e., those lying on some root-leaf path), a process conveniently described by a coloring of the tree edges. In the second step we identify the various colors with distinct unit-length vectors, and map each vertex v to the sum of the vectors corresponding to the colors of the edges along the path from the root to v. We first exhibit a simple constant distortion (but high-dimensional) embedding in Section 2.1; we then show how to reduce the dimensionality in Section 2.2.

2.1 Constant distortion embeddings

Let $T = (V, E)$ be a tree rooted at r with (positive) edge lengths $\ell : E \to \mathbb{R}^+$; by scaling, we can assume that all edge have at least unit length. Let d denote the metric induced on T. An *edge-coloring* of T with colors from $\mathcal{C}$ is simply a map $\chi : E \to \mathcal{C}$. Since we have a rooted tree,

Algorithm BOUNDED-DISTORTION:
(Initially all the edges are uncolored.)
1. set $k = \lceil \log_2 \mathsf{diam}(T) \rceil$ and $Y_0 = \emptyset$.
2. **for** $i = 1, \ldots, k$ **do**
3. $\quad$ set $r_i = 2^{k-i}$ and let $Y_i \supseteq Y_{i-1}$ be r_i-net of T.
4. $\quad$ **for** every $y \in Y_i \setminus Y_{i-1}$ (in an arbitrary order) **do**
5. $\quad\quad$ color all the uncolored edges in P_y with a new color

Figure 2.1. Edge-coloring the tree T

the ancestor-descendant relationship is well-defined: Now a *monotone path* in T is the simple path between a vertex and one of its descendants. We call an edge-coloring *proper* if each color-class forms a monotone path in T. Finally, χ is an α-*good coloring* of T if χ is proper, and for every vertex $v \in V$ and every ancestor u of v, the path from u to v in T contains a monochromatic portion of length at least $\alpha \cdot d(u, v)$. The following lemma is standard, and thus a proof is omitted (see, for instance, [28]).

Lemma 2.1. *Let $T = (V, E)$ be a rooted tree with non-negative edge lengths. If T has an α-good coloring then $c_p(T) \leq \frac{2}{\alpha}$ for all $p \in [1, \infty]$.*

We now define a certain class of "bad comb" metrics, and then show that doubling metrics do not have submetrics which are arbitrarily bad combs. A *k-comb* is a metric induced by an edge-weighted tree T having the following properties: It has a distinguished vertex c called the *center*, and k edge-disjoint simple paths $p_1, \ldots, p_k$ called *hairs*. For a constant γ, the comb is called γ-*bad* if for some value $L \geq \max_i d(c, p_i)$, the length of every hair is in the interval $[L, \gamma L]$. We omit the proof of the next lemma, as it follows easily from Proposition 1.1.

Lemma 2.2. *A 2^γ-bad k-comb (X, d) has doubling constant $\lambda_X > k^{1/(\gamma+3)}$.*

It thus suffices to exhibit good colorings for tree metrics which exclude bad combs.

Theorem 2.3. *Let $0 < \alpha \leq 1/60$. Then every tree metric has either an α-good coloring or a submetric which is a 4-bad $(1/40\alpha)$-comb.*

We prove here the unweighted case of Theorem 2.3; the weighted case is more involved and thus deferred to the full version. (Notice that subdividing edges might increase the doubling constant; consider for instance an n-leaf star with edge lengths 2^i for $i = 1, \ldots, n$.)

Proof of unweighted case. Let P_z be the set of edges in the path from the root to z in the tree T, and let P_{xy} be the edges on the path between x and y. We proceed by showing an algorithm that finds an α-good coloring, unless the tree

metric has a 4-bad $(1/40\alpha)$-comb. The algorithm is given in Figure 2.1.

Algorithm BOUNDED-DISTORTION clearly colors every edge exactly once and thus defines an edge-coloring χ; we claim χ is proper. Indeed, the algorithm maintains the invariant that the colored edges form a connected subtree of T containing the root. It follows that the edges colored in any single execution of step 4 form a monotone path. Since we use a new color each time around, the claim follows.

Now assume that χ is not α-good. Then there exist vertices u, v with u an ancestor of v and $D = d(u, v) = |\mathsf{P}_{uv}|$ with the following property: if $u = w_0, w_1, \ldots, w_\ell = v$ are the *color-transition* vertices on P_{uv}, i.e., the subpath $\mathsf{P}_{w_j w_{j+1}}$ is a maximal monochromatic segment for each j, then $d(w_j, w_{j+1}) < \alpha D$ for all j. For each w_j, consider the execution of Step 5 in which $\mathsf{P}_{w_{j-1} w_j}$ falls in P_y and is colored, and let y_j denote the vertex y at this point in time. (See Figure 2.2.)
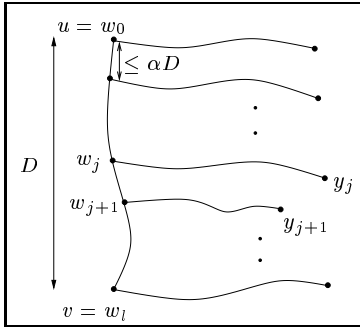

Figure 2.2. An α-bad coloring.

Let m be such that $D/4 < r_m \leq D/2$. We now claim that for any j, if $d(v, w_j) \geq D/2$ then $y_j \in Y_m$. Indeed, Y_m is an $r_m \leq D/2$ net, hence $B(v, D/2)$ must contain a net point $y' \in Y_m$. Furthermore, this net point must be a descendant of w_j in T. It follows that all the edges in the path $\mathsf{P}_{y'}$ are colored by the time y' is considered as a net point. Since that path contains $\mathsf{P}_{w_{j-1} w_j}$, we know that y_j is considered no later than y', and thus $y_j \in Y_m$.

Finally, consider all the vertices w_j with $d(u, w_j) \leq D/10$. We just proved that all the corresponding y_j are in Y_m; since Y_m is an r_m-net, $d(y_j, y_{j+1}) \geq r_m > D/4$. This implies that either $d(w_j, y_j) \geq D/10$ or $d(w_{j+1}, y_{j+1}) \geq D/10$, as otherwise we would contradict the triangle inequality $d(y_j, y_{j+1}) \leq d(y_j, w_j) + d(w_j, w_{j+1}) + d(w_{j+1}, y_{j+1})$. We can now obtain the comb: for each j, at least one of $\mathsf{P}_{w_j y_j}$ and $\mathsf{P}_{w_{j+1} y_{j+1}}$ is a hair of length at least $D/10$. (Since the edges are unit-weighted, we can cut off the paths to a length of exactly $D/10$). These hairs are edge-disjoint because they are colored differently, and the total number of hairs is at least $\frac{1}{2}(\frac{D/10}{\alpha D} - 3) \geq \frac{1}{40\alpha}$. Letting u be the center of the comb, we

have $d(u, w_j) \leq D/10$ for each hair (w_j, y_j), so we indeed obtained a 4-bad $(1/40\alpha)$-comb. (For unit-weight trees we actually get a 1.01-bad comb.) $\qquad\square$

Theorem 2.4. *Every doubling tree metric T has $c_p(T) = O(1)$.*

Proof. Let $\lambda = \lambda_T$ be the doubling constant of the metric T. Setting $\gamma = 2$ in Lemma 2.2, the tree T does not contain a 4-bad λ^6-comb. Let $\alpha = 1/40\lambda^6$; since $\lambda \geq 2$, the parameter α is bounded above by $1/60$, and hence Theorem 2.3 implies that T has an α-good coloring. Finally, applying Lemma 2.1 gives us an embedding with distortion $c_p(T) \leq 80\lambda^6$. $\qquad\square$

2.2 Constant dimension or frugal coloring

The algorithm of Section 2.1 gave us a constant distortion embedding into ℓ_p spaces; however, it used up to a linear number of dimensions. In this section, we reduce this drastically by embedding into a *constant* number of dimensions.

Theorem 2.5. *Every doubling tree metric embeds into $\ell_p^{O(1)}$ with $O(1)$ distortion for every $p \in [1, \infty]$.*

The proof again proceeds by edge-coloring the tree $T = (V, E)$, this time with $O(1)$ colors; of course, the coloring can no longer be proper, and we will have to reanalyze the embedding. We will still make use of the ideas given in Lemmas 2.1 and 2.3. In particular, we edge-color the tree using only colors from a set $\mathcal{C}$ of $|\mathcal{C}| = \lambda^{O(\log \lambda)}$ colors; essentially, we apply algorithm **Bounded-Distortion** with an unbounded number of temporary colors (not from $\mathcal{C}$), and at the end of each iteration we replace the temporary colors by colors from $\mathcal{C}$. We prove here the unweighted case and defer the extension to weighted trees to the full version.

Proof of unweighted case. For a path p, let $\ell(p)$ be the length of p. A coloring $\chi : V \to \mathcal{C}$ will be called α-*reasonable* if, for every pair $u, v \in V$ with $x = \mathsf{lca}(u, v)$, there exists a color $c \in \mathcal{C}$ such that the following holds: If $\mathcal{P}_u$ is the set of paths between x and u which are colored c, and $\mathcal{P}_v$ is the same for v, then

$$\left| \sum_{p \in \mathcal{P}_u} \ell(p) - \sum_{p \in \mathcal{P}_v} \ell(p) \right| \geq \alpha \cdot d(u, v).$$

It is straightforward that applying the embedding of Lemma 2.1 to an α-reasonable coloring of T with k colors yields an embedding of T into ℓ_p^k with distortion at most α. (In fact, the dimension can be reduced further to $O(\frac{1}{\alpha^2} \log k)$ by using vectors β_i that are near-orthogonal; the details are omitted from this version of the paper.) Thus it suffices to show that T has an α-reasonable coloring.

We now color the edges of T using algorithm BOUNDED-D&D given in Figure 2.2. Let λ be the doubling constant of T, and set $\alpha = 1/40\lambda^6$ as in Theorem

2.4. Let $\mathcal{C}$ be a set of $|\mathcal{C}| = \lambda^{3 \log(68/\alpha)} = \lambda^{O(\log \lambda)}$ colors. To complete the proof of Theorem 2.5 it suffices to show that this algorithm produces an $\alpha/2$-reasonable coloring χ of T using only colors from $\mathcal{C}$. To this end, Lemma 2.7 below shows that the algorithm can be implemented with these few colors, and Lemma 2.8 proves that the coloring the algorithm produces is $\alpha/4$-reasonable. $\quad\square$

We will make use of the following proposition.

Proposition 2.6. *Along any root-leaf path, the edges that are colored in iteration i have total length at most $2r_i$. In particular, $\ell(p) \le 2r_i$ for every path $p \in P_i$.*

Proof of Proposition 2.6. We can assume $i > 1$; the claim is trivial for $i = 1$ since $2r_1 > \mathsf{diam}(T)$. For a leaf v, let $S_{i,v}$ be the subset of edges along P_v that are actually colored in iteration i. Consider the edge (x, y) of $S_{i,v}$ farthest from the root, with $y \in Y_i \setminus Y_{i-1}$. Since $y \notin Y_{i-1}$, there exists $y' \in Y_{i-1}$ with $d(y, y') \le r_{i-1} = 2r_i$, and furthermore, $S_{i,v} \subseteq \mathsf{P}_y \setminus \mathsf{P}_{y'}$. It follows that $S_{i,v}$ is contained in the path between y and $\mathsf{lca}(y, y')$, and thus its length is at most $d(y, \mathsf{lca}(y, y')) \le d(y, y') \le 2r_i$. $\quad\square$

Lemma 2.7. *Step 9 of algorithm **Bounded-D&D** can be done greedily with $|\mathcal{C}| = \lambda^{3 \log(68/\alpha)}$.*

Proof of Lemma 2.7. The proof idea is very simple: For each path $p \in P_i$, it suffices to show that the number of paths that are *adjacent* to p and were permanently colored before p is less than $|\mathcal{C}| - 1$, and hence we can color p greedily. We will, in fact, show something slightly stronger: We show that the number of paths $p' \in P_i^*$ with $p \leftrightarrow p'$ is at most $|\mathcal{C}| - 1$. Let $z \in p$ be its endpoint farther from the root. Let $p' \in P_i^*$ be adjacent to p, and let $y' \in p' \cap Y_i$ be the net point that caused p' to be colored (in some iteration no later than i). We now split the paths p' into two types.

Type 1: $\ell(p') \le \frac{17}{\alpha} r_i$. In this case, y' is close to z; quantitatively, $d(z, y') \le \ell(p) + d(p, p') + \ell(p') \le 2r_i + \frac{16}{\alpha} r_i + \frac{17}{\alpha} r_i < \frac{34}{\alpha} r_i$. Hence each such path p' corresponds to a distinct point $y' \in B(z, \frac{34}{\alpha} r_i)$. But y' belongs to the r_i-net Y_i, so Proposition 1.1 implies that the number of such points y', and hence the number of paths p', is at most $\lambda^{\lceil \log 68/\alpha \rceil} \ll |\mathcal{C}|/2$.

Type 2: $\ell(p') > \frac{17}{\alpha} r_i$. These paths p' form hairs of length at least $\frac{17}{\alpha} r_i$ in a comb centered at z. (Since we are in a unit-weighted tree, we can truncate these paths to get the correct length.) The distance of each hair from the center z is $d(z, p') \le \ell(p) + d(p, p') \le 2r_i + \frac{16}{\alpha} r_i \le \frac{17}{\alpha} r_i$. Lemma 2.2 implies that X cannot have a 1.01-bad λ^3-comb, and thus the number of such hairs (and thus the number of paths p') is less than $\lambda^4 \le |\mathcal{C}|/4$.

Hence the total number of paths $p' \in P_i^*$ that are adjacent to any single path $p \in P_i$ is less than $\frac{3}{4}|\mathcal{C}| \le |\mathcal{C}| - 1$, and we can extend the coloring to p. $\quad\square$

Lemma 2.8. *Algorithm **Bounded-D&D** computes an $\alpha/4$-reasonable coloring.*

Proof of Lemma 2.8. Let $u, v \in V$ with $x = \mathsf{lca}(u, v)$, and assume without loss of generality that $t = d(x, u) \ge \frac{1}{2} d(u, v)$. Recall that $\alpha = 1/40\lambda^6 \le 1/60$; hence Theorem 2.4 guarantees a monochromatic path p of length at least αt between x and u. We will show that the lengths of all the paths between x and v with color $\chi(p)$ add up to at most $\alpha t/2$, which immediately implies that χ is $\alpha/4$-reasonable, as desired.

Consider any maximal monochromatic path p' between x and v with $\chi(p) = \chi(p')$, and let i' be the iteration in which p' is colored. We claim that $i' > i_0 + 2$ where i_0 is the maximum index such that $2r_{i_0} \ge \alpha t$. (We may assume that $\alpha t \ge 1$ and thus $r_{i_0} \le \alpha t \le 2r_{i_0}$.) Assuming this claim, let us prove the lemma. For every such value of i', Proposition 2.6 implies that the corresponding paths p' have total length at most $2r_{i'}$; now summing over all values $i' > i_0 + 2$ shows that the total length of paths having color $\chi(p)$ is at most $\sum_{i' > i_0 + 2} 2r_{i'} \le 4r_{i_0+3} = r_{i_0}/2 < \alpha t/2$.

It remains to prove the claim. Assume for contradiction that $i' \le i_0 + 2$. We now have two cases, depending on the iteration i at which p was colored. The first case is when $i' \le i$, and hence $p' \in P_i^*$. Clearly, $d(p, p') \le d(u, v) \le 2t$, and Proposition 2.6 implies that $\alpha t \le 2r_i$; putting the two together gives $d(p, p') \le \frac{4}{\alpha} r_i$. Hence p, p' are adjacent at iteration i and cannot have the same color, which contradicts the assumption that $\chi(p) = \chi(p')$. The second case is where $i < i' \le i_0 + 2$, in which case $p \in P_{i'}^*$. A similar argument applies; since $d(p, p') \le 2t \le \frac{4}{\alpha} r_{i_0} \le \frac{16}{\alpha} r_{i'}$, we have that p, p' are adjacent at iteration i', giving us the desired contradiction. $\quad\square$

3 Clustering

In this section, we give decomposition theorems (for general metrics) whose performance behaves smoothly in terms of their doubling dimension, yielding greatly improved results when the dimension is bounded.

3.1 The padded decomposition

First, we describe a useful low-diameter decomposition for metric spaces. Under other guises, such decompositions are the main tool in many embeddings, and have numerous other applications.

For a metric space (X, d) and a subset $S \subseteq X$, let $\mathsf{diam}(S)$ denote the diameter of the submetric induced on S. Let $\mathcal{P}$ be the collection of all the partitions of X. Given a partition $P \in \mathcal{P}$ and $x \in X$, define

$$\pi_P(x) = \sup\{t : \exists C \in P \text{ with } B(x, t) \subset C\}.$$

Definition 3.1. *An (r, ε)-padded probabilistic decomposition of a metric (X, d) is a distribution μ over $\mathcal{P}$ satisfying:*

Algorithm BOUNDED-D&D: (Initially all the edges are uncolored.)

1. set $k = \lceil \log_2 \mathsf{diam}(T) \rceil$ and $Y_0 = \emptyset$.
2. **for** $i = 1, \ldots, k$ **do**
3.   set $r_i = 2^{k-i}$ and let $Y_i \supseteq Y_{i-1}$ be r_i-net of T.
4.   **for** every point $y \in Y_i \setminus Y_{i-1}$ (in an arbitrary order) **do**
5.   temporarily color all the uncolored edges in P_y with a single new color.
6.   let P_i be the set of temporarily colored paths.
7.   let $P_i^* = P_i \cup \{$all paths colored with a color from $\mathcal{C}\}$.
8.   **remark** paths $p_1, p_2 \in P_i^*$ are *adjacent*, written $p_1 \leftrightarrow p_2$, if $d(p_1, p_2) \leq \frac{16}{\alpha} r_i$.
9.   recolor the (temporarily colored) paths in P_i with colors from $\mathcal{C}$
       such that $\chi(p_1) \neq \chi(p_2)$ for all $p_1, p_2 \in P_i^*$ with $p_1 \leftrightarrow p_2$.

Figure 2.3. Frugal edge-coloring of the tree T

---

1. *Bounded diameter:* $\mathsf{diam}(C) \leq r$ *for every cluster* C *in every partition* P *in the support of* μ.

2. *Padding:* $\Pr_\mu[\pi_P(x) \geq \varepsilon r] \geq \frac{1}{2}$ *for all* $x \in X$.

Such decompositions have been given earlier for general metrics, where $1/\varepsilon = O(\log |X|)$. We show that such decompositions exist where ε depends only on the doubling dimension of X. The probabilistic technique we use is inspired by the analysis of [7].

Theorem 3.2. *Let* (X, d) *be a finite metric space. Then for every* $r > 0$ *there exists an* (r, ε)-*padded probabilistic decomposition of* X *with* $1/\varepsilon \leq 64 \dim(X)$.

Proof. For ease of notation, we construct a $(4r, \varepsilon)$-padded decomposition. Let N be an r-net of X. Let σ be a random permutation on N, and choose a radius R uniformly at random from $(r, 2r)$. For each $y \in N$, define a cluster

$$C_y = \big\{ x \in X : x \in B(y, R) \text{ and } \sigma(y) < \sigma(z)$$
$$\text{for all } z \in N \text{ with } x \in B(z, R) \big\}.$$

Clearly, $\mathsf{diam}(C_y) \leq 4r$. Finally, let $P = \{C_y\}_{y \in N}$ and note it is a partition of X because N is an r-net and $R \geq r$.

Now fix a point $x \in X$ and some $t \in [0, r]$. Let $W = B(x, 2r + t) \cap N$, and note that $m = |W| \leq 8^{\dim(T)}$ by Proposition 1.1. Arrange the points $w_1, \ldots, w_m \in W$ in order of increasing distance from x, and let I_k be the interval $[d(x, w_k) - t, d(x, w_k) + t]$. Let us say that $B(x, t)$ is *cut* by a cluster C_{w_k} if $C_{w_k} \cap B(x, t) \neq \emptyset$ but $B(x, t) \not\subseteq C_{w_k}$. Finally, write $\mathcal{E}_k$ for the event that w_k is the minimal element in W (according to σ) for which $C_{w_k} \cap B(x, t) \neq \emptyset$ and C_{w_k} cuts $B(x, t)$. Then,

$$
\begin{aligned}
\Pr[B(x, t) \text{ is cut}] &\leq \sum_{k=1}^m \Pr[\mathcal{E}_k] \\
&= \sum_{k=1}^m \Pr[R \in I_k] \cdot \Pr[\mathcal{E}_k \mid R \in I_k] \\
&\leq \sum_{k=1}^m \frac{2t}{r} \cdot \frac{1}{k} \leq \frac{2t}{r}(1 + \ln m),
\end{aligned}
$$

and the latter quantity is at most $\frac{8t}{r} \dim(X)$. Finally, note that when $B(x, t)$ is not cut, we have $\pi_P(x) \geq t$. Setting $t = \frac{4r}{64 \dim_C(X)}$, we get $\Pr[\pi_P(x) \geq t] \geq \frac{1}{2}$ as desired. $\square$

3.2 Locality and dimension

Decompositions like that given by Theorem 3.2 have seen numerous applications in recent years. Often, one wants a small distribution, i.e. a multi-set of m partitions $\mathcal{D} = [P_1, \ldots, P_m]$ such that the uniform distribution on $\mathcal{D}$ satisfies condition (2) of Definition 3.1.

Usually, this is accomplished by choosing $m = \Theta(\log n)$ independent partitions according to μ; a Chernoff bound then yields the desired result. Here, we show that $m = O(\dim(X) \log \dim(X))$ partitions suffice. The idea is to use the Lovász Local Lemma along with some ideas borrowed from [18] to exploit the locality of the padded decomposition.

Theorem 3.3. *For a metric space* (X, d), *let* $1/\varepsilon = 512 \dim(X)$. *Then for any* $r \geq 0$, *there exists a multi-set* $\mathcal{D} = [P_1, \ldots, P_m]$ *of* $m = O(\dim(X) \log \dim(X))$ *partitions such that*

1. *For every* $C \in \bigcup_{i=1}^m P_i$, $\mathsf{diam}(C) \leq r$.

2. *If* P *is chosen uniformly from* $\mathcal{D}$, *then for all* $x \in X$,

$$\Pr_{P \in \mathcal{D}}[\pi_P(x) \geq \varepsilon r] \geq \frac{1}{2}.$$

Proof of Theorem 3.3. Let $r > 0$ be fixed and let μ be as in Theorem 3.2 (note that μ depends on r). Let $P_1, \ldots, P_m$ be partitions of X chosen according to μ (for some m to be chosen later). For $x \in X$, let Y_x be the number of P_i for which $\pi_{P_i}(x) \leq 2\varepsilon r$. The probability of the latter event is at most $\frac{1}{8}$ according to the analysis of Theorem 3.2, and thus $\mathbb{E}Y_x \leq \frac{m}{8}$. Finally, let $\mathcal{E}_x^m$ be the event that $Y_x > \frac{m}{2} = 4\,\mathbb{E}Y_x$. A standard Chernoff bound (see e.g. [1]) shows that $\Pr[\mathcal{E}_x^m] \leq (9/10)^m$.

7

Claim 3.4. *If N is an εr-net in X then $\Pr\left[\bigwedge_{y \in N} \overline{\mathcal{E}_y^m}\right] > 0$.*

Notice that the above claim suffices to prove the Theorem. Indeed, if $\pi_P(y) \geq 2\varepsilon r$ for every $y \in N$, then by the triangle inequality, $\pi_P(x) \geq \varepsilon r$ for every $x \in X$.

In proving the claim, we will require the following symmetric form of the Lovász Local Lemma (see, e.g., [1]).

Lemma 3.5 (Lovász Local Lemma). *Let $A_1, \dots, A_n$ be events in an arbitrary probability space. Suppose that each A_i is mutually independent of all but at most d other events A_j, and suppose that $\Pr[A_i] \leq p$ for all $1 \leq i \leq n$. If $ep(d+1) \leq 1$ then $\Pr[\wedge_{i=1}^n \overline{A_i}] > 0$.*

We now claim that $\mathcal{E}_y^m$ is mutually independent of all events $\mathcal{E}_{y'}^m$ for which $d(y, y') > 4r$. To see this, note that every cluster formed under μ (according to Theorem 3.2) has diameter at most r, thus no such cluster can simultaneously cut $B(y, r)$ and $B(y', r)$; the claimed independence follows.

It follows that we can upper bound the number of events that are non-mutually independent of $\mathcal{E}_y^m$ by $d = |B(y, 4r) \cap N| \leq \dim(X)^{O(\dim(X))}$ (by Proposition 1.1). Now if $\Pr[\mathcal{E}_y^m] \leq \frac{1}{e(d+1)}$, then the local lemma implies Claim 3.4. But this is easily accomplished by choosing some $m = O(\dim(X) \log \dim(X))$. $\qquad\square$

4 Upper bounds for doubling metrics

4.1 Bourgain-style embeddings via nets

Here, we give a simple proof that doubling metrics embed into Euclidean space with $O(\sqrt{\log n})$ distortion. This is based on Bourgain's proof for general metrics, with the added twist that we take distances from subsets of *nets*, rather than subsets of the entire metric. A full proof is deferred to the full version.

Theorem 4.1. *For any n-point doubling metric (X, d), the distortion required to embed X into ℓ_2 is $c_2(X) = O(\sqrt{\log n})$.*

Proof (sketch). For every scale $2^k, k \in \mathbb{Z}$, construct a 2^k-net of X and define a map $\varphi_k : X \to \ell_2$ by $\varphi_k(x) = \sqrt{1/2^{|Y_k|}}\, (d(x, A))_{A \in Y_k}$, where Y_k is the set of all subsets of this 2^k-net. The final map is basically the normalized direct sum of all such maps, i.e. $\oplus_k 2^{-\frac{1}{2}|Y_k|}\varphi_k$, though standard considerations must be made to achieve a distortion which depends only on n and not the aspect ratio of X. $\qquad\square$

It turns out that the dependence of the distortion on $\dim(X)$ in the above theorem is exponential, and its worst-case performance is much worse than Bourgain's $O(\log n)$ bound. We will remedy this in the next section, using the decomposition theorems of Section 3.

By adapting the analysis of the above proof we can prove the following theorem. In order to reduce the dimension to $O(\log n)$, we first bound the distortion among points in certain nets using the Lovász Local Lemma, and then use it to deduce a bound on the distortion among all points. The $1 + \epsilon$ distortion is achieved by suitably modifying various constants in the proof.

Theorem 4.2. *For any fixed $\epsilon > 0$, every doubling metric $(1 + \epsilon)$-embeds into $\ell_\infty^{O(\log n)}$.*

4.2 Embeddings via clustering

In this section, we obtain embeddings that degrade gracefully with the doubling dimension; these are based on the decomposition theorems of Section 3. The next theorem uses an embedding technique due to Rao [34]. Our analysis requires considerably more effort since we must keep tight control on the dimension of our embedding (this improves the dimension of the host space by a factor of $\Omega(\log n)$, say, for doubling metrics, and is essential for Section 4.2.1).

Theorem 4.3. *For any metric (X, d) and any $p \in [1, \infty)$,*

$$c_p(X) = O\left(\dim(X) \cdot (\log n)^{\min(\frac{1}{2}, \frac{1}{p})}\right).$$

Proof (outline). Let (X, d) be an arbitrary metric and fix a value $r \geq 0$. We discuss the case $p = 2$; the proof for other values of p is similar. We produce a map $\Phi_r : X \to \mathbb{R}^k$ with $k = O(\dim(X) \log \dim(X))$ as follows. Let $P_1, \dots, P_m$ be the multi-set of decompositions guaranteed by Theorem 3.3. Fix some $i \in \{1, \dots, m\}$, and for every cluster $C \in P_i$, choose a value $\mathrm{sgn}(C) \in \{-1, +1\}$ uniformly at random. Following Rao [34] (see also [29, Ch. 15]), define the map $\varphi_i : X \to \mathbb{R}$ by

$$\varphi_i(x) = \sum_{C \in P_i} \mathrm{sgn}(C) \cdot d(x, X \setminus C).$$

Finally, set $\Phi_r = \frac{1}{\sqrt{m}}(\varphi_1 \oplus \cdots \oplus \varphi_m)$.

Let ε be as in Theorem 3.3, let N_r be an $\frac{\varepsilon r}{64}$-net, and set $S_r = \{\{x, y\} \in N_r^2 : d(x, y) > r\}$. For $\{x, y\} \in S_r$, let $\mathcal{E}_{x,y}$ be the event that $\|\Phi_r(x) - \Phi_r(y)\| < \frac{1}{8}\varepsilon r$. The following lemma is deferred to the full version.

Lemma 4.4. *$\Pr\left[\bigwedge_{\{x,y\} \in S_r} \overline{\mathcal{E}_{x,y}}\right] > 0$. In other words, there exists a map Φ_r such that for all $\{x, y\} \in S_r$, $\|\Phi_r(x) - \Phi_r(y)\| \geq \frac{1}{8}\varepsilon r$.*

For any $i \in \mathbb{Z}$, let $\Phi^{(i)} = \Phi_{2^i}$, and set $\Phi = \bigoplus_{i \in \mathbb{Z}} \Phi^{(i)}$. The proof that Φ satisfies the statement of Theorem 4.3 appears in the full version. $\qquad\square$

4.2.1 Assouad's Theorem

Recall that for $0 < \alpha \leq 1$, (X, d^α) is the metric that arises from a metric (X, d) by raising all distances to the power α. In his seminal paper [2], Assouad proved the following (perhaps surprising) result: For any doubling metric, and every fixed $0 < \alpha < 1$, there exist constants k and D (depending only on the doubling constant of X) such that (X, d^α) embeds into ℓ_2^k with distortion at most D.

Although Assouad's proof is easily converted to an algorithm, the parameters k and D grow exponentially with $\dim(X)$. We present an algorithmic version based on the formation of coordinates in Section 4.2 and a trick of Assouad [2] to avoid interaction between scales. The values k and D that we achieve are both nearly linear in $\dim(X)$.

Theorem 4.5. *Let (X, d) be an arbitrary metric, and fix a value $0 < \alpha < 1$. Then there exist values $k = O(\dim(X) \log \dim(X))$ and $D = O(\dim(X))$ such that (X, d^α) can be embedded into ℓ_2^k with distortion at most D.*

We now show the construction of the embedding. A proof is deferred to the full version. Let $m = m(\alpha)$, choose $\{e_1, \ldots, e_m\}$ to be an orthonormal basis of $\mathbb{R}^m$, and extend it to an infinite periodic sequence $\{e_i\}_{i \in \mathbb{Z}}$. Now simply define $\Phi : X \to \mathbb{R}^k$ by

$$\Phi(x) = \sum_{i \in \mathbb{Z}} 2^{\alpha i} \Phi^{(i)}(x) \otimes e_i.$$

5 Lower bounds on distortion

In this section we show that the dependence on n in the upper bounds of Theorems 4.1 and 4.3 is necessary. In fact, the next theorem shows that for Euclidean embeddings ($p = 2$) of doubling metrics, the aforementioned upper bounds are existentially tight. (It extends easily to a tight lower bound for any fixed $p \geq 2$.)

Theorem 5.1. *There exists a family of metrics (G_k, d_k) which are uniformly doubling and series-parallel, such that $c_2(G_k) = \Omega(\sqrt{\log |G_k|})$.*

We now describe the metrics (G_k, d_k); they are shortest path metrics on weighted series-parallel graphs. G_0 consists of a single edge of weight 1. G_{k+1} is obtained from G_k by replacing every edge in (u, v) in G_k with the six-edge configuration shown in Figure 5.4. The weight of each new edge is 4^{-k}. It is easy to see that (G_k, d_k) is a submetric of (G_{k+1}, d_{k+1}).

In [20, 21] it is shown that a similar family of metrics is uniformly doubling (see [22] for a very simple proof). The lower bound follows from an appropriate Poincaré inequality whose basis is the following geometric fact: For any six points $s, t, a, b, c, d \in L_2$, we have

$$||s - t||^2 + ||b - d||^2 \leq 4(||a - s||^2 + ||c - t||^2)$$
$$+ 2(||a - b||^2 + ||b - c||^2 + ||c - d||^2 + ||d - a||^2).$$

By summing appropriately weighted versions of this inequality over all six-point subsets corresponding to "copies" of G_1, it is possible to obtain a simple Poincaré inequality which immediately yields the desired lower bound.
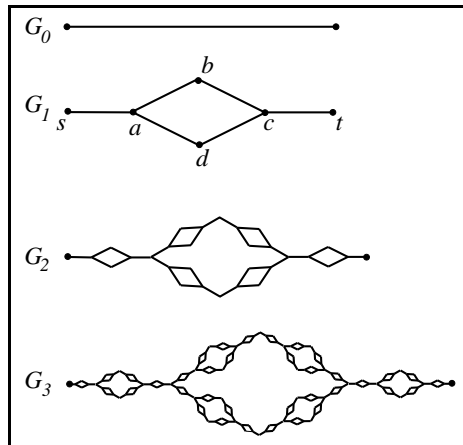

Figure 5.4. The lower bound graphs.

6 Discussion

Applications. We briefly mention several algorithmic applications of our results, deferring details to the final version. Using the decomposition theorems in Section 3, we can show that doubling metrics admit $(k, \sqrt{\log n})$-volume respecting embeddings for all k; doubling metrics can also be embedded into distributions of *doubling* trees (HSTs of bounded degree) with distortion $O(\log n)$, and they embed into the line with constant average distortion. These results can be used to improve approximation algorithms and on-line algorithms for doubling metrics.

The approximation ratio of the minimum bandwidth (linear ordering) problem can be improved using our results to $O(\log^2 n)$ for doubling trees and $O(\log^3 n)$ for general doubling graphs, improving over the general case by $\sqrt{\log n}$ and $\sqrt{\log \log n}$ factors respectively. Furthermore, our tree coloring in conjunction with a modification of [12] show that the bandwidth of a graph with local density β is at most $O(\beta^{1.5} \log^2 n)$.

Distance-labelings of graphs assign labels to vertices so that the distance between two vertices can be computed (approximately) from their labels alone. For simplicity, assume that all pairwise distances are integers and are bounded by Δ. Theorem 4.2 can then be used to give labels for doubling metrics with $O(\log n \log \Delta)$ bits, allowing to approximate the distances within a $(1 + \epsilon)$ factor. This has an exponential dependence on $\dim(X)$, which can be improved to $O(\dim(X) \log \Delta \log n (\frac{4}{\epsilon})^{\dim(X)})$ bits per label for any $0 < \epsilon < 1$, using techniques from Section 3.

Finally, the techniques of Section 3 can be used to exhibit

9

an $O(\dim(T))$-approximation for the 0-extension problem where (T, d) is the metric on terminals.

Dimensionality reduction. In previous sections, we studied the distortion of general metrics in terms of their doubling dimension. Here, let us consider an n-point set $X \subseteq \ell_2^n$ with $\dim(X) = O(1)$. In other words, although X inhabits some high-dimensional space, it looks very low-dimensional in terms of volume. The Johnson-Lindenstrauss flattening lemma [16] tells us that there is a $1 + \epsilon$ embedding of X into $\ell_2^{O(\epsilon^{-2} \log n)}$. We pose the following intriguing question which was asked independently in [22].

Question 1. *Can every doubling submetric of ℓ_2 be $O(1)$-embedded into $\ell_2^{O(1)}$?*

It can be shown that no linear projection can achieve such a result (as in [16]), but there are a number reasons to believe that its resolution might be positive; these are discussed in the full version.

Acknowledgments. We thank Yair Bartal, Manor Mendel, Assaf Naor, and Oded Schramm for some enlightening discussions, and Stephen Semmes for useful pointers to the literature.

References

[1] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience, New York, second edition, 2000.

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

[3] Y. Aumann and Y. Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.*, 27(1):291–301, 1998.

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

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

[6] B. Brinkman and M. Charikar. On the impossibility of dimension reduction in ℓ_1. To appear in 44th Annual IEEE Symposium on Foundations of Computer Science, 2003.

[7] G. Calinescu, H. Karloff, and Y. Rabani. Approximation algorithms for the 0-extension problem. In *12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 8–16. SIAM, 2001.

[8] F. R. K. Chung and P. D. Seymour. Graphs with small bandwidth and cutwidth. *Discrete Math.*, 75(1-3):113–119, 1989.

[9] K. L. Clarkson. Nearest neighbor queries in metric spaces. In *29th Annual ACM Symposium on Theory of Computing*, pages 609–617. ACM, New York, 1997.

[10] M. M. Deza and M. Laurent. *Geometry of cuts and metrics*. Springer-Verlag, Berlin, 1997.

[11] U. Feige. Approximating the bandwidth via volume respecting embeddings. *J. Comput. System Sci.*, 60(3):510–539, 2000.

[12] A. Gupta. Improved bandwidth approximation for trees and chordal graphs. *J. Algorithms*, 40(1):24–36, 2001.

[13] A. Gupta, I. Newman, Y. Rabinovich, and A. Sinclair. Cuts, trees and l_1-embeddings of graphs. In *40th Annual Symposium on Foundations of Computer Science*, pages 399–408. IEEE, 1999.

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

[15] P. Indyk. Algorithmic applications of low-distortion geometric embeddings. In *42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 10–33, 2001.

[16] W. B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

[17] D. Karger and M. Ruhl. Finding nearest neighbors in growth-restricted metrics. In *34th Annual ACM Symposium on the Theory of Computing*, pages 63–66, 2002.

[18] R. Krauthgamer and J. R. Lee. The intrinsic dimensionality of graphs. In *35th Annual ACM Symposium on Theory of Computing*, pages 438–447. ACM, 2003.

[19] R. Krauthgamer and J. R. Lee. Navigating nets: Simple algorithms for proximity search. Manuscript, July 2003.

[20] T. J. Laakso. Ahlfors Q-regular spaces with arbitrary $Q > 1$ admitting weak Poincaré inequality. *Geom. Funct. Anal.*, 10(1):111–123, 2000.

[21] 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.

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

[23] J. R. Lee, M. Mendel, and A. Naor. Metric structures in L_1: Dimension, snowflakes, and average distortion. Manuscript, July 2003.

[24] J. R. Lee and A. Naor. Embedding the diamond graph in L_p and dimension reduction in L_1. Manuscript, June 2003.

[25] N. Linial. Finite metric spaces - combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians III*, pages 573–586, 2002.

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

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

[28] J. Matoušek. On embedding trees into uniformly convex Banach spaces. *Israel J. Math.*, 114:221–237, 1999.

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

[30] I. Newman and Y. Rabinovich. A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.*, 29(1):77–81, 2003.

[31] T. S. E. Ng and H. Zhang. Predicting internet network distance with coordinates-based approaches. In *21st Annual Joint Conference of the IEEE Computer and Communications Society (INFOCOM-02)*, pages 170–179, 2002.

[32] C. G. Plaxton, R. Rajaraman, and A. W. Richa. Accessing nearby copies of replicated objects in a distributed environment. *Theory Comput. Syst.*, 32(3):241–280, 1999.

[33] Y. Rabinovich. On average distortion of embedding metrics into the line and into l1. In *35th Annual ACM Symposium on the Theory of Computing*, pages 456–462, 2003.

[34] S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *15th Annual Symposium on Computational Geometry*, pages 300–306. ACM, 1999.

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

[36] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.