Labelings vs. Embeddings: On Distributed Representations of Distances [*]

Arnold Filtser[†] Lee-Ad Gottlieb[‡] Robert Krauthgamer[§]

Abstract

We investigate for which metric spaces the performance of distance labeling and of ℓ_∞-embeddings differ, and how significant can this difference be. Recall that a distance labeling is a distributed representation of distances in a metric space (X, d), where each point $x \in X$ is assigned a succinct label, such that the distance between any two points $x, y \in X$ can be approximated given only their labels. A highly structured special case is an embedding into ℓ_∞, where each point $x \in X$ is assigned a vector $f(x)$ such that $\|f(x) - f(y)\|_\infty$ is approximately $d(x, y)$. The performance of a distance labeling or an ℓ_∞-embedding is measured via its distortion and its label-size/dimension.

We also study the analogous question for the prioritized versions of these two measures. Here, a priority order $\pi = (x_1, \dots, x_n)$ of the point set X is given, and higher-priority points should have shorter labels. Formally, a distance labeling has prioritized label-size $\alpha(.)$ if every x_j has label size at most $\alpha(j)$. Similarly, an embedding $f : X \to \ell_\infty$ has prioritized dimension $\alpha(\cdot)$ if $f(x_j)$ is non-zero only in the first $\alpha(j)$ coordinates. In addition, we compare these their prioritized measures to their classical (worst-case) versions.

We answer these questions in several scenarios, uncovering a surprisingly diverse range of behaviors. First, in some cases labelings and embeddings have very similar worst-case performance, but in other cases there is a huge disparity. However in the prioritized setting, we most often find a strict separation between the performance of labelings and embeddings. And finally, when comparing the classical and prioritized settings, we find that the worst-case bound for label size often "translates" to a prioritized one, but also a surprising exception to this rule.

1 Introduction

It is often useful to succinctly represent the pairwise distances in a metric space (X, d) in a distributed manner. A common model, called *distance labeling*, assigns to each point $x \in X$ a label $l(x)$, such that some algorithm $\mathcal{A}$ (oblivious to (X, d)) can compute the distance between any two points $x, y \in X$ given only their labels $l(x), l(y)$, i.e., $\mathcal{A}(l(x), l(y)) = d(x, y)$. The goal is to construct a labeling whose label-size, defined as $\max_{x \in X} |l(x)|$, is small. For general n-point metric spaces, Gavoille, Peleg, Pérennes and Raz [GPPR04] constructed a labeling scheme with label size of $O(n)$ words, and also proved that this bound to be tight.[1]

To obtain smaller label size, one often considers algorithms that approximate the distances. A distance labeling is said to have *distortion* $t \geq 1$ if

$$\forall x, y \in X, \qquad d(x, y) \leq \mathcal{A}(l(x), l(y)) \leq t \cdot d(x, y).$$

While the lower bound of [GPPR04] holds even for distortion $t < 3$, Thorup and Zwick [TZ05] constructed a labeling scheme with distortion $2t - 1$ and label size $O(n^{1/t} \log n)$ for every integer $t \geq 2$. These bounds are almost tight (assuming the Erdős girth conjecture), and demonstrate that for distortion $O(\log n)$, label size $O(\log n)$ is possible.

From an algorithmic viewpoint, there is a significant advantage to labels possessing additional structure, for example labels that are vectors in a normed space. This structure can lead to improved algorithms, for example nearest neighbor search [Ind98, BG19]. A natural candidate for vector labels is the ℓ_∞ space, since every finite metric space embeds into it isometrically

[*] This is a preliminary version. A full version of the paper is available at arxiv:1907.06857.

[†] Columbia University. Email: arnold273@gmail.com. Most of the work was conducted while the author was affiliated with Ben-Gurion University of the Negev. Work supported in part by Simons Foundation, ISF grant 1817/17, and by BSF Grant 2015813.

[‡] Ariel University. Email: leead@ariel.ac.il.

[§] Weizmann Institute of Science. Email: robert.krauthgamer@weizmann.ac.il. Work partially supported by ONR Award N00014-18-1-2364, the Israel Science Foundation grant #1086/18, and a Minerva Foundation grant.

[1] We measure size in words to avoid issues of bit representation. In the common scenario where distances are polynomially-bounded integers, every word has $O(\log n)$ bits, where $n = |X|$. The bounds in [GPPR04] are given in bits and are for unweighted graphs. Nevertheless, once we consider weighted graphs, $O(n)$ words are sufficient and necessary for exact distance labeling, see Theorem 2.1.

(i.e., with no distortion). As such isometric embeddings require $\Omega(n)$ dimensions [LLR95], one may consider instead embeddings with small distortion. Formally, an embedding $f : X \to \ell_\infty$ is said to have *distortion* $t \geq 1$ if

$$\forall x, y \in X, \qquad d(x, y) \leq \|f(x) - f(y)\|_\infty \leq t \cdot d(x, y).$$

Matoušek [Mat96] showed that for every integer $t \geq 2$, every metric space embeds with distortion $2t-1$ into ℓ_∞ of dimension $O(n^{1/t} \cdot t \cdot \log n)$ (which again is almost tight assuming the Erdős girth conjecture). For distortion $O(\log n)$, Abraham *et al.* [ABN11] later improved the dimension to $O(\log n)$.

In this paper, we take the perspective that ℓ_∞-embeddings are a particular form of distance labelings, and study the trade-offs these two models offer between distortion and dimension/label-size. While the inherent structure of ℓ_∞-embeddings makes them preferable, one may suspect that their additional structure precludes the tight trade-off achieved using generic labelings. Yet we have seen that for general metric spaces, the performance of ℓ_∞-embeddings is essentially equivalent to that of generic labelings. This observation motivates us to consider more restricted input metrics, such as ℓ_p spaces, planar graph metrics, and trees. The central question we address is the following.

Question 1.1. *In what settings are generic distance labelings more succinct than ℓ_∞-embeddings, and how significant is the gap between them?*

Priorities. Elkin, Filtser and Neiman [EFN18] have introduced the problems of *prioritized distortion* and *prioritized dimension*; they posit that some points have higher importance or priority, and it is desirable that these points achieve improved performance. Formally, given a priority ordering $\pi = \{x_1, \ldots, x_n\}$ on the point set X, we say that embedding $f : X \to \ell_\infty$ possesses *prioritized contractive distortion*[2] $\alpha : \mathbb{N} \to \mathbb{N}$ (w.r.t. π) if for all $j < i$

$$(1.1) \quad \frac{d(x_j, x_i)}{\alpha(j)} \leq \|f(x_j) - f(x_i)\|_\infty \leq d(x_j, x_i)\ .$$

Prioritized distortion is defined similarly for distance labeling. Furthermore, we say that a labeling scheme has prioritized label-size $\beta : \mathbb{N} \to \mathbb{N}$, if every x_j has label length $|l(x_j)| \leq \beta(j)$. We say that embedding $f : X \to \ell_\infty$ has prioritized dimension β if every $f(x_j)$ is non-zero only in the first $\beta(j)$ coordinates (i.e., $f_i(x_j) = 0$ whenever $i > \beta(j)$). Here too ℓ_∞-embeddings are a more structured case of labelings, and we again ask what are the possible trade-offs and how these two compare. It is worth noting that the priority functions α, β are defined on all of $\mathbb{N}$ and apply when embedding every finite metric space; in particular, they are not allowed to depend on $n = |X|$. We ask analogously to Question 1.1 about prioritized label size/dimension.

Question 1.2. *In what settings are distance labelings with prioritized label size more succinct than ℓ_∞ embeddings with prioritized dimension, and how significant is the gap between them?*

In many embedding results, the (worst-case) distortion is a function of the size of the metric space $n = |X|$. Elkin *et al.* [EFN18] demonstrated a general phenomenon: Often a worst-case distortion $\alpha(n)$ can be replaced with a prioritized distortion $\tilde{O}(\alpha(j))$ using the same α. [3] For example, every finite metric space embeds into a distribution over trees with prioritized expected distortion $O(\log j)$, which extends the $O(\log n)$ known from [FRT04]. Recently, Bartal *et al.* [BFN19] showed that every finite metric space embeds into ℓ_2 with prioritized distortion $O(\log j)$, which extends the $O(\log n)$ known from [Bou85]. In fact, we are not aware of any setting where it is impossible to generalize a worst-case distortion guarantee to a prioritized guarantee. The final question we raise is the following.

Question 1.3. *Does this analogy between worst-case and prioritized distortion extend also to dimension and to label-size, or perhaps their worst-case and prioritized versions exhibit a disparity?*

1.1 Results: Old and New

Our main results and most relevant prior bounds are discussed below and summarized in Table 1. Addition related work is described in Section 1.2.

General Metrics. As discussed above, embeddings and labeling schemes for general graphs have essentially the same parameters. But for prioritized labelings and embeddings, the comparison is more complex. For exact labeling, one can obtain label size $O(j)$ by simply storing in the label of x_j its distances to $x_1, \ldots, x_{j-1}$ (recall that we count words). This is essentially optimal, even if we allow distortion up to 3, see Theorem 2.1. In contrast, for embeddings into ℓ_∞, we show in Theorem 2.2 that prioritized dimension is impossible for distortion less than $\frac{3}{2}$. Specifically, we provide an example

[2] In the original definition of prioritized distortion in [EFN18], the requirement of equation (1.1) is replaced by the requirement $d(x_j, x_i) \leq \|f(x_j) - f(x_i)\|_\infty \leq \alpha(j) \cdot d(x_j, x_i)$. We add the word *contractive* to emphasize this difference. Prioritized contractive distortion is somewhat weaker in that it does not imply scaling distortion (see Section 1.2).

[3] We use $\tilde{O}$ notation to suppress constants and logarithmic factors, that is $\tilde{O}(\alpha(j)) = \alpha(j) \cdot \mathrm{polylog}(\alpha(j))$.

Worst-Case Label-Size/Dimension						
		Distortion	**Distance Labeling**		**Embedding into ℓ_∞**	
1.	General Metric	< 3	$\Theta(n)$ ♣	[GPPR04]	$\Theta(n)$	[Mat13]
2.	General Metric	$O(\log n)$	$O(\log n)$	[TZ05]	$\Theta(\log n)$	[ABN11]
3.	ℓ_p for $p \in [1,2]$	$1+\epsilon$	$O(\epsilon^{-2}\log n)$	(Thm. 3.1)	$\Theta(n)$ ♡	(Thm. 3.2)
4.	Tree	1	$O(\log n)$	[TZ01]	$\Theta(\log n)$	[LLR95]
5.	Planar	1	$\Theta(\sqrt{n})$	[GPPR04]	$\Theta(n)$	[LLR95]
6.	Treewidth k	1	$O(k \log n)$	[GPPR04]	$\Theta(n)$ ♠	[LLR95]
Prioritized Label-Size/Dimension						
		Distortion	**Distance Labeling**		**Embedding into ℓ_∞**	
7.	General Metric	$< \frac{3}{2}$	$\Theta(j)$ ♣	(Thm. 2.1)	$\Theta(n)$ ♢	(Thm. 2.2)
8.	General Metric	$O(\log j)$	$O(\log j)$	[EFN18]	$O(j)$	(Cor. 2.1)
9.	ℓ_p for $p \in [1,2]$	$1+\epsilon$	$O(\epsilon^{-2}\log j)$	(Thm. 3.1)	$j^{\Omega(\frac{1}{\epsilon})}$	(Thm. 3.3)
10.	Tree	1	$O(\log j)$	[EFN18]	$\Theta(\log j)$	(Thm. 4.1)
11.	Planar	1	$\Theta(j)$	(Thm. 5.2)	$\Theta(n)$ ♢	(Thm. 5.1)
12.	Treewidth k	1	$O(k \log j)$	[EFN18]	$\Theta(n)$ ♠ ♢	(Thm. 5.1)

Table 1: *Summary of our findings. Question 1.1 is answered by comparing the last two columns of rows 1-6; in the very general and very restricted families (lines 1,2,4), labelings and embeddings perform similarly, while other families (lines 3,5,6) exhibit a strict separation. Question 1.2 is answered by comparing the last two columns of rows 7-12; we see a strict separation between them in all families other than trees (line 10). Question 1.3 is answered by comparing each row $i = 1, \dots, 6$ with row $i + 6$; for distance labeling, worst-case bound $\beta(n)$ translates to prioritized $O(\beta(j))$ except for planar graphs (lines 5,11), while for embeddings, dimension translates to its prioritized version only for trees (lines 4,10).*
Table footnotes: ♣ *The upper bound is for distortion 1 (i.e. isometric embedding).* ♡ *Holds for $1 + \epsilon < \sqrt{2}$ and $p \in [1, \infty]$.* ♠ *Holds for $k \geq 2$.* ♢ *This excludes priority dimension for any function $\alpha : \mathbb{N} \to \mathbb{N}$ that is independent of $n = |X|$.*

where the images of x_1 and x_2 must differ in at least $\Omega(n)$ coordinates for arbitrarily large n. This proves a strong separation between embeddings and labelings, and also demonstrates an embedding result that has no prioritized counterpart.

For prioritized distortion $O(\log j)$, Elkin *et al.* [EFN18] constructed a labeling with prioritized label size of $O(\log j)$. We construct in Theorem 2.3 ℓ_∞-embeddings with different tradeoffs between the prioritized distortion α and dimension β. Two representative examples are prioritized distortion $\alpha(j) = O(\log j)$ with prioritized dimension $\beta(j) = O(j)$, and $\alpha(j) = O(\log\log j)$ with $\beta(j) = O(j^2)$. This is significantly better than for the $O(1)$-distortion case, yet considerably weaker than results on labeling.

Additional interesting results in this context were given in [EFN18], showing that every metric space embeds into every ℓ_p, $p \in [1, \infty]$, with prioritized distortion $O(\log^{4+\epsilon} j)$ and prioritized dimension $O(\log^4 j)$ (for every constant $\epsilon > 0$). Furthermore, Elkin and Neiman [EN19] have communicated to us that they obtained two additional embeddings into ℓ_∞ with: (1) prioritized distortion $2\left\lceil\frac{k\log j}{\log n}\right\rceil - 1$ and prioritized dimension $O(k \cdot n^{\frac{1}{k}} \cdot \log n)$; and (2) prioritized distortion $2k\log\log j+1$ and prioritized dimension $O(k \cdot j^{\frac{2}{k}} \cdot \log n)$. Note that the dimension bounds of [EN19] depend on $n = |X|$ and hence are not truly prioritized. See Table 2 for a comparison of these results with ours.

ℓ_p **Spaces.** The seminal Johnson-Lindenstrauss Lemma [JL84] states that every n-point subset of ℓ_2 embeds with distortion $1 + \epsilon$ into $\ell_2^{O(\epsilon^{-2}\log n)}$ (where as usual ℓ_p^d denotes a d-dimensional ℓ_p space), and this readily implies a labeling with distortion $1+\epsilon$ and label size $O(\epsilon^{-2}\log n)$. Since every ℓ_p, $p \in [1,2]$, embeds isometrically into squared-L_2 (equivalently, its snowflake embeds into L_2), this implies a labeling with the same performance for ℓ_p as well, see Theorem 3.1. Furthermore, we show in Theorem 3.1 (using [NN18]) that this labeling can be prioritized to achieve distortion $1 + \epsilon$ with label size $O(\epsilon^{-2}\log j)$.

For ℓ_∞- embeddings, the performance is significantly worse. We show in Theorem 3.2 that for certain n-point subsets of ℓ_p, for any $p \in [1, \infty]$, embedding into ℓ_∞ with distortion less than $\sqrt{2}$ requires $\Omega(n)$ coordinates. (Recall that $O(n)$ coordinates are sufficient to isometrically embed every n-point metric into ℓ_∞.) For prioritized embeddings into ℓ_∞ with distortion $1+\epsilon$, we prove a lower bound of $j^{\Omega(\frac{1}{\epsilon})}$ on the prioritized dimension, see Theorem 3.3.

Tree Metrics. In their seminal paper on metric embeddings, Linial, London and Rabinovich [LLR95] proved that every n-node tree embeds isometrically into $\ell_\infty^{O(\log n)}$. In the context of routing, Thorup and Zwick [TZ01] constructed an exact distance labeling with label size $O(\log n)$ (where routing decisions can be done in constant time), and Elkin *et al.* [EFN18] modified it to achieve prioritized label size $O(\log j)$. Our contribution is a prioritized version of [LLR95], i.e., an isometric embedding of a tree metric into ℓ_∞ with prioritized dimension $O(\log j)$, see Theorem 4.1. We note that an equivalent result was proved independently and concurrently by Elkin and Neiman [EN19].

Planar Graphs and Restricted Topologies. We first consider exact distance labeling and isometric embeddings. Gavoille *et al.* [GPPR04] showed that planar graphs admit exact labeling with label size $O(\sqrt{n})$, and proved a matching lower bound.[4] They further showed that treewidth-k graphs admit exact labeling with label size $O(k \log n)$. Linial *et al.* [LLR95] proved that an isometric embedding of the n-cycle graph into ℓ_∞, and in fact into any normed space, requires $\Omega(n)$ coordinates.[5] Notice that the cycle graph is both planar and has treewidth 2; hence, there is a strict separation between distance labeling and ℓ_∞-embedding.

For exact distance labeling, we prove that planar graphs require prioritized label size $\Omega(j)$ (based on [GPPR04]), see Theorem 5.2. This bound is tight, as prioritized label size $O(j)$ is possible already for general graphs Theorem 2.1. We conclude that priorities make exact distance labelings much harder for planar graphs.[6] This lower bound for exact prioritized labeling holds for unweighted graphs as well, hence this type of labeling is now well understood. For embedding of treewidth-k graphs, Elkin *et al.* [EFN18] constructed exact labeling with prioritized label size $O(k \log j)$. For isometric embeddings into ℓ_∞, we show in Theorem 5.1 that no prioritized dimension is possible for the cycle graph, which provides a lower bound for both planar and treewidth-2 graphs. This implies a dramatic separation for these families.

[4]This lower bound, as well as all other lower bounds from [GPPR04], count bits rather than words.

[5]Their proof is much more general than what is required for ℓ_∞. For a simpler proof for the special case studied here, see Theorem 5.1.

[6]Interestingly, for unweighted planar graphs, Gavoille *et al.* [GPPR04] prove only a lower bound of $\Omega(n^{\frac{1}{3}})$ on the label size, and closing the gap to the upper bound $O(\sqrt{n})$ remains an important open question.

Additional results on labelings with $1+\epsilon$ distortion, and embeddings with constant distortion are described in Section 1.2.

Conclusions. We uncover a wide spectrum of settings and bounds that answer our questions. For Question 1.1, in the simplest case of trees, labeling and embeddings have similar behavior, and both admit prioritization with similar bounds. For the least restricted case of general graphs/metrics, we find similarly that labelings and embeddings exhibit similar behavior across various distortion parameters. However between these two extremes, for ℓ_p spaces, planar graphs and treewidth k graphs, we see significant separations between labelings and embeddings.

For Question 1.2, we show that labelings admit far superior prioritized versions than their embedding counterparts in all settings other than trees, and most notably for general graphs and for planar/bounded-treewidth graphs, where no prioritized dimension is possible. In ℓ_p spaces, while we did not ruled out the possibility of prioritized dimension, we demonstrate a surprising exponential gap between labelings and embeddings (also in the dependence on ϵ).

For Question 1.3 we saw that labeling schemes have prioritized versions, in all cases other than planar graphs where instead of the desired $O(\sqrt{j})$ label size we show that $\Theta(j)$ is surprisingly necessary. For embeddings into ℓ_∞ we showed that for larger distortion some prioritized dimension is possible, even though it is much worse that its labeling counterpart.

Some interesting remaining open questions are presented in Section 6.

1.2 Related Work

For distortion $1 + \epsilon$ in planar graphs, Thorup [Tho04] and Klein [Kle02] constructed distance labels of size $O(\log n/\epsilon)$. Abraham and Gavoille [AG06] generalized this result to K_r-minor-free graphs, achieving label size $O(g(r) \log n/\epsilon)$. [7] No low-dimension embedding into ℓ_∞ with distortion $1 + \epsilon$ is known for planar graphs or even treewidth-2 graphs. If one allows larger distortion, Krauthgamer *et al.* [KLMN05] proved that planar graphs embed with distortion $O(1)$ into $\ell_\infty^{O(\log n)}$, or more generally that K_r-minor-free graphs embed with distortion $O(r^2)$ into $\ell_\infty^{O(3^r \cdot \log n)}$. Abraham *et al.* [AFGN18] showed that K_r-minor-free graphs embed with distortion $O(1)$ into $\ell_\infty^{O(g(r) \log^2 n)}$. Turning to priorities, Elkin *et al.* [EFN18] constructed prioritized versions of distance labeling with distortion $1 + \epsilon$. Specifically, for planar and K_r-minor-free graphs they achieve label sizes of $O(\log j/\epsilon)$ and

[7]The function $g(r)$ depends only on r and is taken from the structure theorem of Robertson and Seymour [RS03].

$O(g(r)\log n/\epsilon)$, respectively. No prioritized embeddings are known, nor lower bounds thereof.

Elkin *et al.* [EFN17] studied the problem of *terminal* distortion, where there is specified a subset $K \subset X$ of terminal points, and the goal is to embed the entire space (X,d) while preserving pairwise distances among $K \times V$. Embeddings with terminal distortion can be used used to construct embeddings with prioritized distortion. We utilize this approach in Theorems 3.1 and 4.1.

Abraham *et al.* [ABN11] studied *scaling* distortion, which provides improved distortion for $1-\epsilon$ fractions of the pairs, simultaneously for all $\epsilon \in (0,1)$, as a function of ϵ. A stronger version called *coarse scaling* distortion has improved distortion guarantees for the farthest pairs. Bartal *et al.* [BFN19] showed that scaling distortion and prioritized distortion (in the sense of [EFN18]) are essentially equivalent, but this is not known to hold for the prioritized contractive distortion we use in the current paper (see footnote (2)).

Another way to represent distances is a distance oracle [TZ05]. This is a data structure that, given a pair of points, returns an estimate of their pairwise distance. The properties of interest are distortion, space and query time. A distance labeling can be viewed as a distributed version of a distance oracles, see also [MN07, Che14, Che15, EFN18, ACE+18, CGMW19].

Exact distance labelings were studied in the precession of bits (i.e. not asymptotically) see e.g. [AGHP16]. Another type of labeling studied is adjacency labeling [AKTZ19], where given two labels one can compute whether the vertices are adjacent. Efficiency of the labeling algorithms has also been studied [WP11].

1.3 Preliminaries The ℓ_p-norm of a vector $x = (x_1,\dots,x_d) \in \mathbb{R}^d$ is $\|x\|_p := (\sum_{i=1}^d |x_i|^p)^{1/p}$, where $\|x\|_\infty := \max_i |x_i|$. An embedding f between two metric spaces (X,d_X) and (Y,d_Y) has distortion $c\cdot t$ if for every $x,y\in X$, $\frac{1}{c}\cdot d_X(x,y) \le d_Y(f(x),f(y)) \le t\cdot d_X(x,y)$. t (resp. c) is the *expansion* (resp. *contraction*) of f. If the expansion is 1, we say that f is *Lipschitz*.

Embedding $f : X \to \ell_\infty^d$ can be viewed as a collection of embeddings $\{f_i\}_{i=1}^d$ to the line $\mathbb{R}$. By scaling we can assume that the embedding is not contractive. That is, if f has distortion t then for every $x,y\in X$ and i, $|f_i(x)-f_i(y)| \le t\cdot d_X(x,y)$ and there is some index $i_{x,y}$ such that $d_X(x,y) \ge |f_{i_{x,y}}(x) - f_{i_{x,y}}(y)|$. We say that the pair x,y is *satisfied* by the coordinate $i_{x,y}$.

We consider connected undirected graphs $G = (V,E)$ with edge weights $w: E \to \mathbb{R}_{>0}$. Let d_G denote the shortest path metric in G. For a vertex $x\in V$ and a set $A \subseteq V$, let $d_G(x,A) := \min_{a\in A} d(x,a)$, where $d_G(x,\emptyset) := \infty$. We often abuse notation and write the graph G instead of its vertex set V.

We always measure the size of a label by the number of words needed to store it (where each word contains $O(\log n)$ bits). For ease of presentation, we will ignore issues of representation and bit counting. In particular, we will assume that every pairwise distance can be represented in a single word. We note however that the lower bounds of [GPPR04] are given in bits, and therefore our Theorem 5.2 is as well.

All logarithms are in base 2. Given a set A, $\binom{A}{2} = \{\{x,y\} \mid x,y\in A, x\neq y\}$ denotes all the subsets of size 2. The notation $x = (1\pm\epsilon)\cdot y$ means $(1-\epsilon)y \le x \le (1+\epsilon)y$.

2 General Graphs

In this section we discuss our result on succinct representations of general metric spaces. We start with the regime of small distortion. Recall that there exist both exact distance labelings with $O(n)$ label size [GPPR04] as well as isometric embeddings into ℓ_∞^n [Mat13], and both results are essentially tight (even if one allows distortion < 3). In the following theorem we provide a prioritized version of the exact distance labeling.

THEOREM 2.1. *Given an n-point metric space (X,d) and priority ordering $X = \{x_1,\dots,x_n\}$, there is an exact labeling scheme with prioritized label size j. This is asymptotically tight, that is every exact labeling scheme must have prioritized label size $\Omega(j)$. Furthermore, for $t<3$, every labeling scheme with distortion t must have prioritized label size $\tilde{\Omega}(j)$.*

Proof. [Proof of Theorem 2.1] We begin by constructing the labeling scheme. The label of x_j simply consists of the index j and $d(x_1,x_j), d(x_2,x_j),\dots,d(x_{j-1},x_j)$. The size bound and algorithm for answering queries are straightforward. If one allows distortion $t < 3$, [GPPR04] proved that every labeling scheme with distortion t must have label size of $\Omega(n)$ bits, or $\tilde{\Omega}(n)$ words. As some vertex must have a label of size $\tilde{\Omega}(n)$, the prioritized lower bound $\tilde{\Omega}(j)$ follows.

Finally, we prove the $\Omega(j)$ lower bound for exact distance labeling. We begin by arguing that some label must be of length $\Omega(n)$ (in words), and then the $\Omega(j)$ lower bound for prioritized label size follows. The proof follows the steps of [GPPR04]. Consider a complete graph with $\binom{n}{2}$ edges all having integer weights in $\{n+1, n+2, \dots, 2n\}$. Note that there are $n^{\binom{n}{2}}$ such graphs, where each choice of weights defines a different shortest path metric. Given an exact labeling scheme, the labels $l(x_1),\dots,l(x_n)$ precisely encode the graph. Following arguments from [GPPR04], the sum of lengths of the labels must be at least logarithmic in the number of different graphs. Thus

$$\max_i |l(x_i)| \geq \frac{1}{n} \cdot \log n^{\binom{n}{2}} = \Omega(n \log n) \ .$$

We conclude that some label length must be of $\Omega(n \log n)$ bits, or $\Omega(n)$ words. $\square$

While under the standard worst-case model distance labelings and embeddings into ℓ_∞ behave identically, we show that the prioritized versions are very different. In the following theorem we show that no prioritized dimension is possible, even if one allows distortion $< \frac{3}{2}$.

Theorem 2.2. *There is no function $\alpha : \mathbb{N} \to \mathbb{N}$ such that every metric space can be embedded into ℓ_∞ with prioritized dimension α and distortion $t < \frac{3}{2}$ (for any fixed t).*

Proof. Consider the family $\mathcal{G}$ of unweighted bipartite graphs $G = (V = L \cup R, E)$ where $|L| = |R| = n$, for large enough n. We first argue that there is a graph $G \in \mathcal{G}$ with the following properties:

(1) For every $u, v \in R$ or $u, v \in L$, we have $d_G(u, v) = 2$.

(2) Every embedding $f : G \to \ell_\infty$ with distortion $2t$ requires $\Omega(n)$ coordinates.

The existence of G follows by a counting argument similar to [Mat13]. Note that $|\mathcal{G}| = 2^{n^2}$. Denote by $\mathcal{G}' \subseteq \mathcal{G}$ the graphs in $\mathcal{G}$ fulfilling property (1). Our first step is to lower bound $|\mathcal{G}'|$. Sample uniformly a graph $G \in \mathcal{G}$. For $u, v \in R$ (resp. $u, v \in L$) let $I_{u,v}$ be an indicator for the event $d_G(u, v) \neq 2$. $I_{u,v}$ occurs if and only if u and v do not have a common neighbor in L (resp. R). Then $\Pr[I_{u,v}] = (\frac{3}{4})^n$. By a union bound, the probability that property (1) does not hold is at most $2 \cdot \binom{n}{2} \cdot (\frac{3}{4})^n$. We conclude that $|\mathcal{G}'| \geq 2^{n^2} \cdot \left(1 - 2 \cdot \binom{n}{2} \cdot (\frac{3}{4})^n\right) \geq \frac{1}{2} \cdot 2^{n^2}$. Matoušek [Mat13] (proposition 3.3.1) implicitly proved that for any subset $\mathcal{G}'$ of $\mathcal{G}$, if all of $\mathcal{G}'$ embeds into ℓ_∞^d with distortion $2t < 3$, then

$$c^{d \cdot n} \geq |\mathcal{G}'| \ ,$$

where $c > 1$ is a constant depending on $3 - 2t$. Thus $d = \Omega(n)$. We conclude that there is a graph $G \in \mathcal{G}$ fulfilling both properties (1), (2).

Consider such a graph $G = (V = L \cup R, E)$. Note that property (1) implies that there are no isolated vertexes, and moreover for every $u \in R$, $v \in L$, $d_G(u, v) \in \{1, 3\}$. Let G' be the graph G along with two new vertices l, r where l (resp. r) has edges to all vertices in R (resp. L). Note that for every $u, v \in V$, $d_G(u, v) = d_{G'}(u, v)$. Set $L' = L \cup \{l\}$ and $R' = R \cup \{r\}$.

Claim 2.1. *Every embedding $f : G' \to \ell_\infty$ with distortion t has $\Omega(n)$ coordinates i for which $f_i(l) \neq f_i(r)$.*

The proof of the claim appears bellow. We conclude that there are $\Omega(n)$ coordinates where at least one of l, r is not mapped to 0. Set π to be any priority ordering wherein l and r have priorities 1 and 2 respectively. For every priority function $\alpha : \mathbb{N} \to \mathbb{N}$, by taking $n \gg \alpha(2), \alpha(1)$, there is no embedding with prioritized dimension α with respect to π. The theorem follows. $\square$

Proof. [Proof of Claim 2.1] We assume that the embedding has expansion at most t, and for every pair of vertices there is a coordinate where the pair is satisfied (i.e. not contracted). Set $\mathcal{A}_i = \left\{\{u, v\} \in \binom{L' \cup R'}{2} \mid d_G(u, v) = i\right\}$ to be all the vertex pairs at distance exactly i. Note that $\binom{L' \cup R'}{2} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$. To satisfy all the pairs in $\binom{L' \cup R'}{2}$, $\Omega(n)$ coordinates are required (this is property (2)). We will show that we can satisfy all the pairs in $\mathcal{A}_1 \cup \mathcal{A}_2$ using $O(\log n)$ coordinates only. Thus satisfying all the pairs in $\mathcal{A}_3$ requires $\Omega(n)$ coordinates.

The clique K_n can be embed isometrically into $\ell_\infty^{\lceil \log n \rceil}$ [LLR95]. Such an embedding can be constructed by simply mapping K_n to different combinations of $\{0, 1\}^{\lceil \log n \rceil}$. As 1 is the minimal distance, we can just embed all the the $2n + 2$ points as a clique using $O(\log(n))$ coordinates. By doing so, all the pairs in $\mathcal{A}_1$ will be satisfied. $\mathcal{A}_2$ equals $\binom{L'}{2} \cup \binom{R'}{2}$. Note that the metric induced on $\binom{L'}{2}$ is just a scaled clique. Thus we can embed all of L' to the vectors $\{\pm 1\}^{O(\log n)}$. Additionally send all of R' to $\vec{0}$. Note that by doing so we satisfied all the pairs in $\binom{L'}{2}$ while incurring no expansion. Similarly we can satisfy all the pairs in $\binom{R'}{2}$ using $O(\log n)$ additional coordinates.

Next we argue that in a coordinate $f : G' \to \mathbb{R}$ where $f_i(l) = f_i(r)$, no pair of $\mathcal{A}_3$ is satisfied. Indeed, every vertex $v \in L' \cup R'$ is at distance 1 from either l or r. As we have expansion at most t, in a coordinate i where $f_i(l) = f_i(r)$ the maximal distance between a pair of vertices v, u is $2 \cdot t$. In particular, for every $\{v, u\} \in \mathcal{A}_3$, $|f_i(x) - f_i(y)| \leq 2 \cdot t < 3$. Thus no pair $\{v, u\} \in \mathcal{A}_3$ is satisfied.

As there must be $\Omega(n)$ coordinates where some pair from $\mathcal{A}_3$ is satisfied, necessarily there are $\Omega(n)$ coordinates where $f_i(l) \neq f_i(r)$. $\square$

Considering that for distortion less than $\frac{3}{2}$ no prioritized dimension is possible, it is natural to ask for what distortion are prioritized embeddings possible? Some previous results of this nature are described in the introduction to [EFN18, EN19]. As exact distance labeling is possible using $O(j)$ labels, it is also natural to ask what distortion can be obtain with prioritized dimension $O(j)$? The following is a meta theorem constructing various trade-offs. We present some specific implications in Corollary 2.1. A comparison between our results and others appears in Table 2.

Consider a monotone function $\beta : \mathbb{N} \to \mathbb{N}$. For $j \in \mathbb{N}$, let $\chi_\beta(j)$ be the minimal i such that $\beta(\chi_\beta(j)) \geq j$.

THEOREM 2.3. *Given a metric space (X, d) with priority ordering $X = \{x_1, \ldots, x_n\}$ and a function $\beta : \mathbb{N} \to \mathbb{N}$, there is an embedding $f : X \to \ell_\infty$ with contractive prioritized dimension $\beta(\chi_\beta(j))$ and prioritized distortion $2 \cdot \chi_\beta(j)$.*

Proof. We suggest that while inspecting the proof, it may be helpful for the reader to focus on the setting $\beta(i) = 2^i$, wherein $\chi_\beta(j) = \lceil \log j \rceil$. Set $S_0 = \emptyset$ and $S_i = \{x_j \mid j \leq \beta(i)\}$. We define embedding f by setting its j'th coordinate to be

$$f_j(x) := d(x, S_{\chi_\beta(j)-1} \cup \{x_j\}) \ .$$

Note that for every j' such that $\chi_\beta(j') > \chi_\beta(j)$, $f_{j'}(j) = 0$. Note also that there may be many points j' with $j' < j$ and yet $f_j(x_{j'}) \neq 0$. Thus x_j is non-zero only in the first $\beta(\chi_\beta(j))$ coordinates as required.

Next we argue the prioritized distortion. It is clear that f is Lipschitz. Consider a pair of vertices x_j, y. Set $\Delta = d(x_j, y)$, and $\alpha_i = d\left(\{x_j, y\}, S_i\right)$. Then $\infty = \alpha_0 > \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_{\chi_\beta(j)} = 0$. There must be some index i such that $\alpha_i - \alpha_{i+1} \geq \frac{\Delta}{2\chi_\beta(j)}$ and $\alpha_{i+1} \leq \frac{\Delta}{2} - \frac{\Delta}{2\chi_\beta(j)}$. Otherwise, as $\alpha_{\chi_\beta(j)} = 0$, by induction $\alpha_{\chi_\beta(j)-q} < \frac{q\Delta}{2\chi_\beta(j)}$. In particular $\alpha_1 < \frac{(\chi_\beta(j)-1)\Delta}{2\chi_\beta(j)} = \frac{\Delta}{2} - \frac{\Delta}{2\chi_\beta(j)}$ and $\infty = \alpha_0 < \frac{\Delta}{2}$, a contradiction.

Choose $z \in S_{i+1}$ such that $d(\{x_j, y\}, z) = \alpha_{i+1}$. Without loss of generality, $d(x_j, z) = d(\{x_j, y\}, z) = \alpha_{i+1}$. In particular, $d(y, z) \geq d(x_j, y) - d(x_j, z) > \frac{\Delta}{2}$. It holds that $d\left(y, S_i \cup \{z\}\right) = \min\left\{d(y, S_i), d(y, z)\right\} \geq \min\left\{\alpha_i, \frac{\Delta}{2}\right\}$. Thus

$$\begin{aligned} \|f(x_j) - f(y)\|_\infty &\geq \left|d(x_j, S_i \cup \{z\}) - d\left(y, S_i \cup \{z\}\right)\right| \\ &= \left|\alpha_{i+1} - \min\left\{\alpha_i, \frac{\Delta}{2}\right\}\right| \geq \frac{\Delta}{2\chi_\beta(j)} \ . \end{aligned}$$

Prioritized distortion $2 \cdot \chi_\beta(j)$ follows. ◻

COROLLARY 2.1. *Given a metric space (X, d) with priority ordering $X = \{x_1, \ldots, x_n\}$,*

1. *For every $t \in \mathbb{N}$, there is an embedding $f : X \to \ell_\infty$ with prioritized distortion $2 \cdot \lceil \frac{\log j}{t} \rceil$ and prioritized dimension $2^t \cdot j$.*

2. *There is an embedding $f : X \to \ell_\infty$ with prioritized distortion $2 \cdot \lceil \log \log j \rceil$ and prioritized dimension j^2.*

Proof. The first case follow by choosing the function $\beta(i) = 2^{t \cdot i}$. Here $\chi_\beta(j) = \lceil \log_{2^t} j \rceil = \lceil \frac{\log j}{t} \rceil$, thus the prioritized distortion is $2 \cdot \lceil \frac{\log j}{t} \rceil$ while the prioritized dimension is $\beta(\chi_\beta(j)) = 2^{t \cdot \lceil \frac{\log j}{t} \rceil} < 2^{t + \log j} = 2^t \cdot j$.

For the second case choose $\beta(i) = 2^{2^i}$. Here $\chi_\beta(j) = \lceil \log \log j \rceil$, thus the prioritized distortion is $2 \cdot \lceil \log \log j \rceil$ and the prioritized dimension is $\beta(\chi_\beta(j)) = 2^{2^{\lceil \log \log j \rceil}} < 2^{2 \cdot 2^{\log \log j}} = \left(2^{2^{\log \log j}}\right)^2 = j^2$. ◻

Note that the first case implies prioritized distortion $2 \cdot \lceil \log j \rceil$ and prioritized dimension $2j$.

3 ℓ_p Spaces

In this section we consider representations of ℓ_p spaces. As these spaces are somewhat restricted, we focus on the $1 + \epsilon$ distortion regime. We begin with the upper bound for distance labeling.

THEOREM 3.1. *For every $\epsilon > 0$, $p \in [1, 2]$ and n points in ℓ_p, there is an $1 + \epsilon$-labeling scheme with label size $O(\epsilon^{-2} \log n)$. Furthermore, given a priority ordering π, there is an $1 + \epsilon$-labeling scheme with prioritized label size $O(\epsilon^{-2} \log j)$.*

Proof. [Proof of Theorem 3.1] We begin by constructing a labeling scheme for a set X on n points in ℓ_2. Then we will generalize the result to ℓ_p for $p \in [1, 2]$.

As a consequence of the Johnson Lindenstrauss lemma [JL84], there is an embedding $f : X \to \ell_2^{O(\epsilon^{-2} \log n)}$ with $1 + \epsilon$ distortion. By simply storing $f(x)$ as the label of $x \in X$, we obtain an $1 + \epsilon$ labeling scheme with $O(\epsilon^{-2} \log n)$ label size.

Next we consider a set X with priority ordering $\pi = \{x_1, x_2, \ldots, x_n\}$. Narayanan and Nelson [NN18] (improving a previous result by Mahabadi *et al.* [MMMR18]) constructed a terminal version of the JL transform: Specifically, given a set K of k points in ℓ_2 there is an embedding f of the entire ℓ_2 space into $\ell_2^{O(\epsilon^{-2} \log k)}$ such that for every $x \in K$ and $y \in \ell_2$, $\|f(x) - f(y)\|_2 = (1 \pm \epsilon)\|x - y\|_2$.

For $i = 0, 1, \ldots \lceil \log \log n \rceil$, set $S_i = \{x_j \mid j \leq 2^{2^i}\}$. Let $f_i : X \to \ell_2^{O(\log |S_i|)}$ be a terminal JL transform w.r.t. S_i. The label of x_j will consist of $f_0(x_j), f_1(x_j), \ldots, f_{\lceil \log \log j \rceil}(x_j)$. Given a query on $x_j, x_{j'}$, where $j < j'$, our answer will be $\|f_{\lceil \log \log j \rceil}(x_j) -$

Embeddings of General Metrics		
Prioritized Distortion	**Prioritized Dimension**	**Notes**
$O(\log^{4+\epsilon} j)$	$O(\log^4 j)$ [EFN18]	$\forall$ constant ϵ.
$2 \cdot \left\lceil \frac{k \log j}{\log n} \right\rceil - 1$	$O(k \cdot n^{1/k} \cdot \log n)$ [EN19]	$\forall k \in \mathbb{N}$
$2k \cdot \log\log j + 1$	$O(k \cdot j^{2/k} \cdot \log n)$ [EN19]	$\forall k \in \mathbb{N}$
$2 \cdot \lceil \log j \rceil$	$2j$ Corollary 2.1	
$2 \cdot \lceil \log\log j \rceil$	j^2 Corollary 2.1	

Table 2: ℓ_∞-*embeddings of general metrics with different trade-offs between prioritized distortion and dimension. Note that the results from [EN19] depend on* n *and hence are not truly prioritized.*

$f_{\lceil \log\log j \rceil}(x_{j'})\|_2$. The distortion follows as $x_j \in S_{\lceil \log\log j \rceil}$. The length of the label of x_j is bounded by

$$\begin{aligned}\sum_{i=0}^{\lceil \log\log j \rceil} O(\epsilon^{-2} \log |S_i|) &= O(\epsilon^{-2}) \cdot \sum_{i=0}^{\lceil \log\log j \rceil} 2^i, \\ &= O(\epsilon^{-2}) \cdot 2^{\lceil \log\log j \rceil + 1} \\ &= O(\epsilon^{-2} \cdot \log j)\ .\end{aligned}$$

words, as required.

To generalize the labeling schemes to ℓ_p for $p \in [1,2]$, we note that every ℓ_p, $p \in [1,2]$, embeds isometrically into squared-L_2, or equivalently, the snowflake of ℓ_p embeds into L_2 (see e.g. [DL97]). Then a labeling scheme for ℓ_2 implies the same performance for ℓ_p as well, the only change being that the computed distances must be squared. $\square$

Next we turn our attention to lower bounds. Every n-point set in ℓ_2 embeds isometrically into any other ℓ_p space, for $p \in [1,\infty]$ (see e.g. [Mat13]). This implies that any lower bound that we prove for ℓ_2 will holds as well for any other ℓ_p space (as the hard example will reside in ℓ_p as well).

THEOREM 3.2. *For every* $p \in [1,\infty]$ *and* $n \in \mathbb{N}$, *there is a set* A *of* $2n$ *points in* ℓ_p, *such that every embedding of* A *into* ℓ_∞ *with distortion smaller than* $2^{\max\{\frac{1}{2}, 1-\frac{1}{p}\}}$ *has dimension at least* n.

Proof. [Proof of Theorem 3.2] Set $A = \{e_1, -e_1, e_2, -e_2, \ldots, e_n, -e_n\}$, the standard basis and its antipodal points (here $\{e_i, -e_i\}$ is an *antipodal* pair). Fix p, and we will prove that every embedding of $A \subseteq \ell_p$ with distortion smaller that $2^{1-\frac{1}{p}}$ into ℓ_∞ requires at least n coordinates. As mentioned above, the lower bound for $p = 2$ holds for all ℓ_p as well; thus the theorem will follow.

We argue that each coordinate can satisfy at most a single antipodal pair. As there are n such pairs, the lower bound follows. Consider a single coordinate $f : A \to \mathbb{R}$. Assume by way of contradiction that there are $e_i, -e_i, e_j, -e_j \in A$ ($i \neq j$) such that $2 \leq |f(e_i) - f(-e_i)|, |f(e_j) - f(-e_j)|$. As $f(e_i), f(-e_i), f(e_j), f(-e_j) \in \mathbb{R}$, by case analysis there must be a pair consisting of one point from $\{f(e_i), f(-e_i)\}$, and one point from $\{f(e_j), f(-e_j)\}$ at distance at least $\min\{|f(e_i) - f(-e_i)|, |f(e_j) - f(-e_j)|\} \geq 2$. But the actual distance between this pair is only $2^{\frac{1}{p}}$. Thus f has distortion $\frac{2}{2^{1/p}} = 2^{1-\frac{1}{p}}$, a contradiction. $\square$

Note that Theorem 3.2 implies a lower bound of $\Omega(j)$ on the prioritized dimension of an embedding from ℓ_p into ℓ_∞, with distortion smaller than $\sqrt{2}$. However, for distortion $1+\epsilon$ we prove a stronger lower bound with exponential dependency on ϵ.

THEOREM 3.3. *For every* $\epsilon \in (0,1)$ *and* $p \in [1,\infty]$ *there is a set of points in* ℓ_p *and a priority ordering, such that every embedding of them into* ℓ_∞ *with distortion* $1+\epsilon$ *has prioritized dimension at least* $j^{\frac{1}{6\epsilon}}$.

Proof. As above, we may assume that $p = 2$. Furthermore, we will assume that $\epsilon < \frac{1}{6}$, as otherwise a better lower bound follows from Theorem 3.2. Let n be large enough, and $H_n = \{\pm 1\}^n \subseteq \ell_2^n$ be the Hamming cube. We start by creating a symmetric subset $A \subset H_n$ (i.e. $A = -A$), such that all the points in A differ in more than $\epsilon' n$ coordinates, for $\epsilon' = 3\epsilon$. The set A is created in a greedy manner. Initially set $S = H_n$ and $A = \emptyset$. First pick an arbitrary pair $x, -x \in S$ from S and add them to A. Delete from S all the points that differ in fewer than $\epsilon' \cdot n$ coordinates from either x or $-x$. Note that when $y \in S$ is deleted, so is its *antipodal* point $-y$. Thus, both S, A are maintained to be symmetric. We continue with this process until S is empty. The number of points that differ by at most $\epsilon' n$ coordinates from every point $v \in H$ is $\sum_{i=0}^{\epsilon' n} \binom{n}{i} \leq \binom{n}{\epsilon' n}(1 + \frac{\epsilon' n}{n - 2\epsilon' n + 1}) < 2\binom{n}{\epsilon' n}$. Therefore for each added vertex we deleted fewer than $2\binom{n}{\epsilon' n}$ points. We conclude that the size of A is lower bounded by

$$|A| \geq \frac{2^n}{2 \cdot \binom{n}{\epsilon' n}} \geq \frac{1}{2} \cdot \frac{2^n}{\left(\frac{en}{\epsilon' n}\right)^{\epsilon' n}} = \frac{1}{2} \cdot 2^{\left(1-\epsilon' \log \frac{e}{\epsilon'}\right)n} > 2 \cdot 2^{\frac{n}{2}} . \tag{3.2}$$

We argue that an embedding f of A into $\mathbb{R}$ can satisfy at most a single antipodal pair $x, -x$. Indeed, assume by way of contradiction that there is $f : A \to \mathbb{R}$ and $x, y \in A$ such that $\sqrt{4n} \leq |f(x) - f(-x)|, |f(y) - f(-y)| \leq (1+\epsilon)\sqrt{4n}$. Similar to the proof of Theorem 3.2, by case analysis, there must be a pair $z \in \{x, -x\}$ and $w \in \{y, -y\}$ such that $|f(z) - f(w)| \geq \min\{|f(x) - f(-x)|, |f(y) - f(-y)|\} \geq \sqrt{4n}$. As both $x, -x$ differs from both $y, -y$ by more that $\epsilon' n$ coordinates, z coincides with w in at least $\epsilon' n$ coordinates. In particular $\|z - w\|_2 \leq \sqrt{(1-\epsilon') \cdot 4n}$. Thus f has distortion at least $\frac{|f(z)-f(w)|}{\|z-w\|_2} \geq \frac{\sqrt{4n}}{\sqrt{(1-\epsilon')\cdot 4n}} > 1+\epsilon$, a contradiction.

Next, let $Y = \{\pm 1\}^{\epsilon' n}\{0\}^{(1-\epsilon')n}$ be the set of all points that attain values $\{\pm 1\}$ in the first $\epsilon' n$ coordinates, and with all other coordinates 0. Consider a coordinate $f : X \to \mathbb{R}$ that sends all of Y to $\vec{0}$. We argue that f will not satisfy any antipodal pair in A. Indeed, consider an antipodal pair $x, -x$. Let $y \in Y$ be the point agreeing with x on the first $\epsilon' n$ coordinates and 0 everywhere else. It holds that

$$\begin{aligned}
&|f(x) - f(-x)| \\
&\leq |f(x) - f(y)| + |f(y) - f(-y)| + |f(-y) - f(-x)| \\
&\leq (1+\epsilon)\left(\|x - y\|_2 + 0 + \|(-x) - (-y)\|_2\right) \\
&= (1+\epsilon) \cdot 2 \cdot \sqrt{(1-\epsilon')\, n} < \sqrt{4n} .
\end{aligned}$$

As each coordinate can satisfy at most a single antipodal pair from A, we conclude that every $1+\epsilon$ embedding of X into ℓ_∞ must be non-zero on Y in at least $|A|/2$ coordinates.

We can now conclude the proof: Assume by way of contradiction that for any set in ℓ_2 there is a $1+\epsilon$ embedding into ℓ_∞ with prioritized dimension $j^{\frac{1}{6\epsilon}}$. Set priority π for X with the points in Y occupying the first $|Y|$ places. By our assumption, there is an $1+\epsilon$ embedding where the points of Y are non-zero only in the first

$$|Y|^{\frac{1}{6\epsilon}} = \left(2^{\epsilon' n}\right)^{\frac{1}{2\epsilon'}} = 2^{\frac{n}{2}} \overset{(3.2)}{<} \frac{|A|}{2}$$

coordinates. Thus the embedding cannot satisfy all the pairs in A, a contradiction. □

4 Trees

In this section, we present an embedding of trees into ℓ_∞ with prioritized dimension $O(\log j)$. We begin by sketching the classic isometric embedding of trees into $\ell_\infty^{O(\log n)}$ due to [LLR95]: First, identify a separator vertex s, such that a split of tree T at s results in the creation of two trees T_1, T_2, each containing at most $\frac{2}{3}n + 1$ vertices, where $T_1 \cap T_2 = \{s\}$. Now create a new coordinate wherein each vertex $v \in T_1$ assumes value $d(v, s)$, while each vertex $x \in T_2$ assumed value $-d(x, s)$. This coordinate satisfies all pairwise distances $T_1 \times T_2$. Recursively (and separately) embed T_1 and T_2 into ℓ_∞, recalling that each has its own copy of s. The two embeddings are then merged by translating T_2 so that its copy of s is mapped to the same vector assumed by the copy of s in T_1.

Given a priority ordering on the vertices $v_1, v_2, \ldots, v_n$, our goal is to create an isometric embedding into ℓ_∞ with prioritized dimension $O(\log j)$. A natural first step would be to devise a terminal embedding: Given terminal set K, embed T into $\ell_\infty^{O(\log |K|)}$ while preserving all pairwise distances $K \times V$. A terminal embedding can be constructed following the lines of the classic embedding by modifying the separator decision rule, and ensuring that after $O(\log |K|)$ recursive steps all terminals are found in different subtree. However, a terminal embedding of this type is too weak to yield a prioritized embedding, since the mapping of all terminals into $\vec{0}$ (subsequent to their first $O(\log k)$ non-zero coordinates) interferes with the distances between non-terminal pairs.

To circumvent this problem, we shall "fold" the terminals one above the other, until ultimately all terminals will fall on a single representative vertex (see Lemma 4.1). During such a folding, some of the non-terminal vertices will fold upon each other as well, but our terminal embedding will be sufficiently robust to ensure that their distances are retained. We will then use this result on terminal embeddings of trees into ℓ_∞ (Lemma 4.1) to derive the stronger result, priority embeddings of trees into ℓ_∞ (Theorem 4.1).

4.1 Terminal Lemma

LEMMA 4.1. *Given a weighted tree $T = (V, E, w)$ and a set K of k terminals, there exist a pair of embeddings $f : T \to \ell_\infty^{O(\log k)}$ and $g : T \to \mathcal{T}$ (into another weighted tree $\mathcal{T}$) such that the following properties hold:*

1. *Lipschitz: For every $x, y \in V$, $\|f(x) - f(y)\|_\infty \leq d_T(x, y)$ and $d_{\mathcal{T}}(g(x), g(y)) \leq d_T(x, y)$.*
2. *Preservation: For every $x, y \in V$, either $\|f(x) - f(y)\|_\infty = d_T(x, y)$ or $d_{\mathcal{T}}(g(x), g(y)) = d_T(x, y)$, or both.*
3. *Terminal Collapse: g maps all of K into a single vertex, i.e. $|g(K)| = 1$.*

Proof. We may assume that all terminals of K are leafs, as otherwise we can simply add a dummy vertex in place of each terminal, and connect the terminal to the dummy vertex with an edge of weight 0. The proof is by induction on k.

Base cases. For the case $k = 1$ we can just return the tree as is, along with the null embedding into ℓ_∞. Next we prove the case of $k = 2$. Denote the two terminals by t_1, t_2, and let P be the unique path in T connecting t_1, t_2. Let $c \in V$ be the midpoint of t_1 and t_2, such that $d_T(t_1, c) = d_T(t_2, c)$. (If c does not exist in V, then add c to V, and split the corresponding middle edge into two new edges joined at c.) Now "fold" P around c. That is, create a new tree $\mathcal{T}$, where path P is replaced by a new path that ends at c, and every $x \in P$ is found on the new path at distance exactly $d_T(x, c)$ from c. Any pair of points in P equidistant from c are merged – and in particular t_1 and t_2 are now the same point, which is the other endpoint of the new path. All the other edges and vertices remain the same. As a result, we obtain an embedding $g : d_T \to \mathcal{T}$ (see Figure 1 for an illustration). It is clear that g is Lipschitz, and moreover $|g(\{t_1, t_2\})| = 1$.

Having specified the function g, we now describe the function f: separate T into two trees T_1,T_2 where $T_1 \cap T_2 = \{c\}$. Set the function $f : V \to \mathbb{R}$ as follows.

$$f(v) = \begin{cases} d_T(v,c) & v \in T_1 \\ -d_T(v,c) & v \in T_2 \setminus \{c\} \end{cases} \quad . \tag{4.3}$$

See Figure 1 for an illustration of function f. We argue that f is Lipschitz: Consider a pair of vertices u, v. If $u, v \in T_i$ (for some i), then by the triangle inequality $|f(u) - f(v)| = |d_T(u, c) - d_T(v, c)| \leq d_T(u, v)$. Otherwise, assume without loss of generality that $u \in T_1$ while $v \in T_2$. The shortest path from u to v must pass through c, thus

$$|f(u) - f(v)| = |d_T(u,c) + d_T(v,c)| = d_T(u,v) \ . \tag{4.4}$$

It remains only to prove the second property (preservation). Consider a pair of vertices u, v. If $u \in T_1$ and $v \in T_2$, then by equation (4.4) $|f(u) - f(v)| = d_T(u, v)$. Otherwise, if $u, v \in T_i$, the shortest path between u and v in T is isomorphic to the shortest path in $\mathcal{T}$, and so $d_{\mathcal{T}}(u, v) = d_T(u, v)$ as required.

Induction step. For $k > 2$ terminals, we will assume by induction that for every tree with $k' < k$ terminals there are embeddings f, g as required above, such that f uses at most $a \cdot \log k'$ coordinates, for $a = \frac{2}{\log(3/2)}$. Consider a tree T, and a terminal set K of size k. Let $s \in V$ be a separator vertex, such that T can be separated into two trees T_1,T_2 where $T_1 \cap T_2 = \{s\}$, and each T_i contains at most $\frac{2}{3}k$ terminals. As all the terminals are leafs, $s \notin K$. Create a single new coordinate $h^s : V \to \mathbb{R}$ defined as follows

$$h^s(x) = \begin{cases} d_T(x,s) & x \in T_1 \\ -d_T(x,s) & x \in T_2 \end{cases} \quad .$$

It is clear that h^s is Lipschitz, and that for every $x \in T_1, y \in T_2$, $|h^s(x) - h^s(y)| = d_T(x, y)$. For $i \in \{1, 2\}$, invoke the induction hypothesis on T_i with terminal set $K_i = T_i \cap K$, creating embedding pair $f_i : T \to \ell_\infty^{a \cdot \log(|K_i|)}$ and $g_i : T \to \mathcal{T}_i$ which together satisfy requirements (1)-(3). By padding with 0-valued coordinates, we can assume that both f_1 and f_2 use exactly $a \cdot \log \frac{2}{3}k$ coordinates. Moreover, by translation we can assume that $f_1(s) = f_2(s) = \vec{0}$. Set f' to be the combined function of f_1, f_2:

$$f'(x) = \begin{cases} f_1(x) & x \in T_1 \\ f_2(x) & x \in T_2 \end{cases} \quad .$$

We argue that the function f' is Lipschitz: For $x, y \in T_i$, $\|f'(x) - f'(y)\|_\infty = \|f_i(x) - f_i(y)\|_\infty \leq d_{T_i}(x, y) = d_T(x, y)$. On the other hand for $x \in T_1$ any $y \in T_2$, using the triangle inequality

$$\begin{aligned} &\|f'(x) - f'(y)\|_\infty \\ &\quad \leq \|f'(x) - f'(s)\|_\infty + \|f'(s) - f'(y)\|_\infty \\ &\quad \leq d_{T_1}(x,s) + d_{T_2}(s,y) \\ &\quad = d_T(x,s) + d_T(s,y) = d_T(x,y) \ . \end{aligned}$$

Set $\tilde{f}$ to be the concatenation of f' with h^s, and it is clear that $\tilde{f}$ is Lipschitz as well. This completes the description of the embedding into ℓ_∞.

For the embedding into the tree, let $\tilde{\mathcal{T}}$ be composed of the trees $\mathcal{T}_1$ and $\mathcal{T}_2$ glued together in $g_1(s), g_2(s)$. Similarly define $\tilde{g} : T \to \tilde{\mathcal{T}}$ as follows

$$\tilde{g}(x) = \begin{cases} g_1(x) & x \in T_1 \\ g_2(x) & x \in T_2 \end{cases} \quad .$$

Using triangle inequality in the same manner as for f', it is clear that $\tilde{g}$ is Lipschitz.

We argue that requirement (2) holds w.r.t. $\tilde{f}, \tilde{g}$. Indeed, for u, v in T_i,

$$\begin{aligned} &\max\left\{\|\tilde{f}(x) - \tilde{f}(y)\|_\infty, d_{\tilde{\mathcal{T}}}(\tilde{g}(x), \tilde{g}(y))\right\} \\ &\quad \geq \max\left\{\|f_i(x) - f_i(y)\|_\infty, d_{T_i}(g_i(x), g_i(y))\right\} \\ &\quad = d_{T_i}(x,y) = d_T(x,y) \end{aligned}$$

On the other hand, for $v \in T_1,\ u \in T_2$,

$$\begin{aligned} &\max\left\{\|\tilde{f}(v) - \tilde{f}(u)\|_\infty, d_{\tilde{\mathcal{T}}}(\tilde{g}(v), \tilde{g}(u))\right\} \\ &\quad \geq |h^s(v) - h^s(u)| = d_T(v,u) \ . \end{aligned}$$

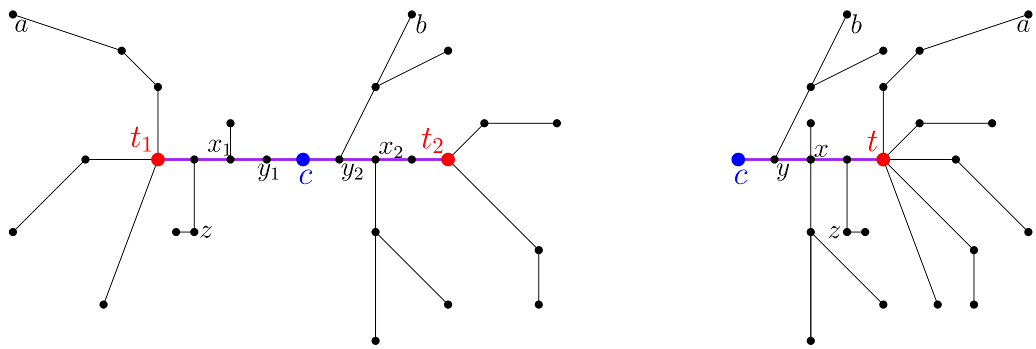

Figure 1: *On the left is illustrated the tree T with two terminals t_1,t_2. The path P between the terminals is colored in purple. The (possibly imaginary) vertex c lies at the midpoint of t_1 and t_2. On the right is illustrated the tree $\mathcal{T}$ which is obtained by "folding" the path P around c.*
In this example, all the edges in T are of unit weight, except for the edge $\{y_1, y_2\}$ that has weight 2. Possible values for the function $f : T \to \mathbb{R}$ (eq. (4.3)) are: $f(t_1) = 4, f(t_2) = -4, f(a) = 7, f(b) = -3, f(x_1) = 2, f(x_2) = -2, f(z) = 4$.

However, requirement (3) does not immediately hold, as $\tilde{\mathcal{T}}$ contains two terminals $g_1(K_1), g_2(K_2)$. Invoke the lemma for the case of $k = 2$ to create two embeddings $\hat{f} : \tilde{\mathcal{T}} \to \mathbb{R}$, $\hat{g} : \tilde{\mathcal{T}} \to \mathcal{T}$ that fulfill requirements (1)−(3). Set $f = \tilde{f} \oplus \hat{f}(\tilde{g})$ to be the concatenation of $\tilde{f}$ with $\hat{f}(\tilde{g})$ and $g = \hat{g}(\tilde{g})$ to be the composition of $\hat{g}$ with $\tilde{g}$ ending in the tree $\mathcal{T}$. It is clear that both f, g are Lipschitz as the Lipschitz property is preserved under concatenation and composition. Moreover, g maps all terminals to a single vertex. Requirement (3) also holds:

$$\begin{aligned} d_T(u,v) &= \max\left\{\|\tilde{f}(v) - \tilde{f}(u)\|_\infty, d_{\tilde{\mathcal{T}}}(\tilde{g}(v), \tilde{g}(u))\right\} \\ &= \max\Big\{\|\tilde{f}(v) - \tilde{f}(u)\|_\infty, \left|\hat{f}(\tilde{g}(v)) - \hat{f}(\tilde{g}(u))\right| \\ &\qquad\qquad , d_{\mathcal{T}}\left(\hat{g}(\tilde{g}(v)), \hat{g}(\tilde{g}(v))\right)\Big\} \\ &= \max\left\{\|f(v) - f(u)\|_\infty, d_{\mathcal{T}}(g(v), g(v))\right\} \ . \end{aligned}$$

Finally, and recalling that $a = \frac{2}{\log(3/2)}$, the number of coordinates is bounded by

$$a \cdot \log \frac{2}{3} k + 1 + 1 = a \cdot \log k + \left(a \cdot \log \frac{2}{3} + 2\right) = a \cdot \log k \ ,$$

The lemma follows. ☐

4.2 Prioritized Embedding of Trees into ℓ_∞

THEOREM 4.1. *Given a weighted tree $T = (V, K, w)$ and a priority ordering π over V, there is an isometric embedding f into ℓ_∞ with prioritized dimension $O(\log j)$.*

Proof. Let $\pi = \{x_1, x_2, \dots, x_n\}$ be a priority order. Set $S_i = \left\{x_i \mid i \le 2^{2^i}\right\}$ for $1 \le i \le \lceil \log\log n \rceil$. Using Lemma 4.1, w.r.t terminal set S_1 construct embeddings $f_1 : T \to \ell_\infty^{O(\log|S_1|)}$ and $g_1 : T \to T_1$. It holds that $g_1(S_1)$ is a single vertex in T_1, and for every $u, v \in V$, $d_T(u,v) = \max\{\|f_1(u) - f_1(v)\|_\infty, d_{T_1}(g_1(u), g_1(v))\}$. Next, using Lemma 4.1 again, w.r.t terminal set $g_1(S_2)$, construct embeddings $f_2 : g_1(T) \to \ell_\infty^{O(\log|S_2|)}$ and $g_2 : g_1(T) \to T_2$. By translation, we can assume that $f_2(g_1(S_1)) = \vec{0}$. Furthermore, $g_2(g_1(S_2))$ is a single vertex in T_2. It also holds that,

$$\begin{aligned} d_T(u,v) = &\quad \max\big\{\, \|f_1(u) - f_1(v)\|_\infty \\ &, \|f_2(g_1(u)) - f_2(g_1(v))\|_\infty, d_{T_2}(g_2(g_1(u)), g_2(g_1(v)))\,\big\} \ . \end{aligned}$$

Generally, in the i step, we invoke Lemma 4.1 on T_{i-1} (w.r.t. terminal set $g_{i-1}(g_{i-2}(\cdots(g_1(S_i))))$) to construct tree T_i and embeddings f_i, g_i. By induction, we constructed trees $T_1, \dots, T_i$ and embeddings $f_1 : T \to \ell_\infty^{O(\log|S_1|)}, \dots, f_i : T_{i-1} \to \ell_\infty^{O(\log|S_i|)}$, $g_1 : T \to T_1, \dots, g_i : T_{i-1} \to T_i$ such that for all $q \in [1, i]$, $g_q(g_{q-1}(\dots(g_1(S_q))))$ is a single vertex in T_q and $f_q(g_{q-1}(\dots(g_1(S_{q-1})))) = \{\vec{0}\}$. Furthermore

(4.5)
$$\begin{aligned} &d_T(u,v) = \\ &\max\big\{\, \|f_1(u) - f_1(v)\|_\infty, \dots \\ &\ , \ \|f_i(g_{i-1}(\dots(g_1(u)))) - f_i(g_{i-1}(\dots(g_1(u))))\|_\infty \\ &\ , \ d_{T_i}\big(g_i(g_{i-1}(\dots(g_1(u)))) \ , \ g_i(g_{i-1}(\dots(g_1(u))))\big)\big\} \ . \end{aligned}$$

Denote $\alpha = \lceil \log\log n \rceil$. After α steps we get functions and trees as above. Set $f = f_1 \oplus (f_2 \circ g_1) \oplus (f_3 \circ g_2 \circ g_1) \oplus \cdots \oplus (f_\alpha \circ g_{\alpha-1} \circ \cdots \circ g_1) : T \to \ell_\infty$. We argue that f is an isomorphism with prioritized dimension $O(\log j)$ as promised. Note that all vertices of V

belong to S_α and hence mapped by $g_\alpha(g_{\alpha-1}(\cdots(g_1)))$ to the same vertex. Thus for every $u,v\in V$, $d_{T_\alpha}\left(g_\alpha(g_{\alpha-1}(\cdots(g_1(u)))), g_\alpha(g_{\alpha-1}(\cdots(g_1(v))))\right) = 0$. By equation (4.5) we get

$$\begin{aligned} d_T(u,v) &= \max\left\{ \|f_1(u)-f_1(v)\|_\infty, \dots \right. \\ &\quad \left., \|f_\alpha(g_{\alpha-1}(\dots(g_1(u)))) - f_\alpha(g_{\alpha-1}(\dots(g_1(u))))\|_\infty \right\} \\ &= \|f(u)-f(v)\|_\infty \end{aligned}$$

Finally we argue that f has prioritized dimension $O(\log j)$. Consider $x_j \in S_{\lceil \log\log j\rceil}$. For every $i > \lceil \log\log j\rceil$ it holds that $f_i\left(g_{i-1}\left(g_{i-2}\left(\cdots(g_1(x_j))\right)\right)\right) = \vec{0}$ (as $x_j \in S_{i-1}$). Therefore x_j might be non-zero only in the first

$$\begin{aligned} \sum_{i=1}^{\lceil \log\log j\rceil} O\left(\log|S_i|\right) &= O\left(\sum_{i=1}^{\lceil \log\log j\rceil} 2^i\right) \\ &= O\left(2^{\lceil \log\log j\rceil+1}\right) = O\left(\log j\right) \end{aligned}$$

coordinates. $\square$

5 Planar Graphs

The theorem below demonstrates that any isometric embedding of the cycle graph C_{2n} into ℓ_∞ requires dimension n. Furthermore, no prioritized dimension is possible for isometric embeddings of the cycle graph. The cycle graph is an interesting example as it is both planar and has treewidth 2. Both proofs of Theorem 5.1 [8] and Theorem 5.2 are omitted from this version.

THEOREM 5.1. *For every $n \in N$, every isometric embedding of C_{2n} (the unweighted cycle graph) into ℓ_∞ requires at least n coordinates. Furthermore, there is no function $\alpha : \mathbb{N} \to \mathbb{N}$ for which the family of cycle graphs $\{C_n\}_{n\in\mathbb{N}}$ can be embedded into ℓ_∞ with prioritized dimension α.*

THEOREM 5.2. *Every prioritized labeling scheme for planar graphs must have prioritized label size of at least $\Omega(j)$ (in bits). This lower bound holds even for unweighted planar graphs.*

6 Open Questions

1. How many coordinates are required in order to embed planar graphs – or even treewidth 2 graphs – into ℓ_∞ with distortion $1+\epsilon$?

2. What is the required label size for $1+\epsilon$ distance labeling for ℓ_p spaces, for $p > 2$?

3. Is it possible to embed ℓ_p spaces ($p \in [1,\infty]$) into ℓ_∞ with distortion $1+\epsilon$ and some prioritized dimension? Theorem 3.3 provided a $j^{\Omega(\frac{1}{\epsilon})}$ lower bound, but did not rule out this possibility. The same question applies when considering constant distortion.

4. All results on embedding of general graphs into ℓ_∞ with both prioritized distortion and dimension (our Theorem 2.2, Theorem 15 in [EFN18], and Theorems 2,3 in [EN19]) feature prioritized *contractive* distortion. What is possible w.r.t. classic prioritized distortion (see footnote (2))?

[8]Theorem 5.1 is a special case of a theorem proved in [LLR95], which applies to general norms. Nonetheless, our proof (which appears in the full version) is much simpler.

References

[ABN11] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. *Advances in Mathematics*, 228(6):3026 – 3126, 2011. 2, 3, 5

[ACE+18] Ittai Abraham, Shiri Chechik, Michael Elkin, Arnold Filtser, and Ofer Neiman. Ramsey spanning trees and their applications. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018*, pages 1650–1664, 2018. 5

[AFGN18] Ittai Abraham, Arnold Filtser, Anupam Gupta, and Ofer Neiman. Metric embedding via shortest path decompositions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018*, pages 952–963, 2018. Full version available at `https://arxiv.org/abs/1708.04073`. 4

[AG06] Ittai Abraham and Cyril Gavoille. Object location using path separators. In *Proceedings of the Twenty-Fifth Annual ACM Symposium on Principles of Distributed Computing, PODC 2006*, pages 188–197, 2006. 4

[AGHP16] Stephen Alstrup, Inge Li Gørtz, Esben Bistrup Halvorsen, and Ely Porat. Distance labeling schemes for trees. In *43rd International Colloquium on Automata, Languages, and Programming, ICALP 2016*, pages 132:1–132:16, 2016. 5

[AKTZ19] Stephen Alstrup, Haim Kaplan, Mikkel Thorup, and Uri Zwick. Adjacency labeling schemes and induced-universal graphs. *SIAM J. Discrete Math.*, 33(1):116–137, 2019. 5

[BFN19] Yair Bartal, Arnold Filtser, and Ofer Neiman. On notions of distortion and an almost minimum spanning tree with constant average distortion. *Journal of Computer and System Sciences*, 2019. Preliminary version appeared in SODA16. 2, 5

[BG19] Yair Bartal and Lee-Ad Gottlieb. Approximate nearest neighbor search for ℓ_p-spaces ($2 < p < \infty$) via embeddings. *Theoretical Computer Science*, pages 27–35, 2019. 1

[Bou85] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52(1-2):46–52, 1985. 2

[CGMW19] Panagiotis Charalampopoulos, Pawel Gawrychowski, Shay Mozes, and Oren Weimann. Almost optimal distance oracles for planar graphs. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019*, pages 138–151, 2019. 5

[Che14] Shiri Chechik. Approximate distance oracles with constant query time. In *Symposium on Theory of Computing, STOC 2014*, pages 654–663, 2014. 5

[Che15] Shiri Chechik. Approximate distance oracles with improved bounds. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015*, pages 1–10, 2015. 5

[DL97] M. M. Deza and M. Laurent. *Geometry of cuts and metrics*. Springer-Verlag, Berlin, 1997. 8

[EFN17] Michael Elkin, Arnold Filtser, and Ofer Neiman. Terminal embeddings. *Theor. Comput. Sci.*, 697:1–36, 2017. 5

[EFN18] Michael Elkin, Arnold Filtser, and Ofer Neiman. Prioritized metric structures and embedding. *SIAM J. Comput.*, 47(3):829–858, 2018. 2, 3, 4, 5, 6, 8, 12

[EN19] Michael Elkin and Ofer Neiman. Lossless prioritized embeddings. *CoRR*, abs/1907.06983, 2019. To appear in SODA 2020. 3, 4, 6, 8, 12

[FRT04] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. Syst. Sci.*, 69(3):485–497, 2004. 2

[GPPR04] Cyril Gavoille, David Peleg, Stéphane Pérennes, and Ran Raz. Distance labeling in graphs. *J. Algorithms*, 53(1):85–112, 2004. 1, 3, 4, 5

[Ind98] P. Indyk. On approximate nearest neighbors in non-Euclidean spaces. In *FOCS*, pages 148–155, 1998. 1

[JL84] William Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984. 3, 7

[Kle02] Philip N. Klein. Preprocessing an undirected planar network to enable fast approximate distance queries. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 820–827, 2002. 4

[KLMN05] R. Krauthgamer, J. R. Lee, M. Mendel, and A. Naor. Measured descent: A new embedding method for finite metrics. *Geometric And Functional Analysis*, 15(4):839–858, 2005. 4

[LLR95] Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995. 2, 3, 4, 6, 9, 12

[Mat96] Jiří Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel Journal of Mathematics*, 93(1):333–344, 1996. 2

[Mat13] Jirı Matoušek. Lecture notes on metric embeddings. Technical report, Technical report, ETH Zürich, 2013. https://kam.mff.cuni.cz/ matousek/ba-a4.pdf. 3, 5, 6, 8

[MMMR18] Sepideh Mahabadi, Konstantin Makarychev, Yury Makarychev, and Ilya P. Razenshteyn. Nonlinear dimension reduction via outer bi-lipschitz extensions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018*, pages 1088–1101, 2018. 7

[MN07] Manor Mendel and Assaf Naor. Ramsey partitions and proximity data structures. *Journal of the European Mathematical Society*, 9(2):253–275, 2007. preliminary version appeared in FOCS06, see `http://portal.acm.org/citation.cfm?id=1170136.1170493`. 5

[NN18] Shyam Narayanan and Jelani Nelson. Optimal terminal dimensionality reduction in euclidean space. *CoRR*, abs/1810.09250, 2018. 3, 7

[RS03] Neil Robertson and Paul D. Seymour. Graph minors. XVI. excluding a non-planar graph. *J. Comb. Theory, Ser. B*, 89(1):43–76, 2003. 4

[Tho04] Mikkel Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *J. ACM*, 51(6):993–1024, 2004. 4

[TZ01] Mikkel Thorup and Uri Zwick. Compact routing schemes. In *SPAA*, pages 1–10, 2001. 3, 4

[TZ05] Mikkel Thorup and Uri Zwick. Approximate distance oracles. *J. ACM*, 52(1):1–24, 2005. 1, 3, 5

[WP11] Oren Weimann and David Peleg. A note on exact distance labeling. *Inf. Process. Lett.*, 111(14):671–673, 2011. 5