Lipschitz Decompositions of Finite ℓ_p Metrics

Robert Krauthgamer
Weizmann Institute of Science, Rehovot, Israel

Nir Petruschka
Weizmann Institute of Science, Rehovot, Israel

Abstract

Lipschitz decomposition is a useful tool in the design of efficient algorithms involving metric spaces. While many bounds are known for different families of finite metrics, the optimal parameters for n-point subsets of ℓ_p, for $p > 2$, remained open, see e.g. [Naor, SODA 2017]. We make significant progress on this question and establish the bound $\beta = O(\log^{1-1/p} n)$. Building on prior work, we demonstrate applications of this result to two problems, high-dimensional geometric spanners and distance labeling schemes. In addition, we sharpen a related decomposition bound for $1 < p < 2$, due to Filtser and Neiman [Algorithmica 2022].

2012 ACM Subject Classification Theory of computation → Computational geometry; Theory of computation → Sparsification and spanners

Keywords and phrases Lipschitz decompositions, metric embeddings, geometric spanners

Digital Object Identifier 10.4230/LIPIcs.SoCG.2025.66

Related Version *arXiv Version*: `https://arxiv.org/abs/2502.01120`

Funding *Robert Krauthgamer*: Work partially supported by the Israel Science Foundation grant #1336/23, by the Israeli Council for Higher Education (CHE) via the Weizmann Data Science Research Center, and by a research grant from the Estate of Harry Schutzman.

Acknowledgements We thank Arnold Filtser and Ofer Neiman for helpful discussions.

1 Introduction

The pursuit of approximating metric spaces by simpler structures has inspired the development of fundamental concepts, such as graph spanners [47, 46] and low-distortion embeddings into various spaces [37, 7], both of which have a wide range of algorithmic applications. Many of these results, including for instance [7, 48, 17, 30, 6, 19], rely on various notions of decomposition of a metric space into low-diameter clusters, and these decompositions are most often randomized. One extensively studied notion, see e.g. [13, 24, 17, 25], is *Lipschitz decomposition* (also called *separating decomposition*), which informally is a random partition of a metric space into low-diameter clusters, with a guarantee that nearby points are likely to belong to the same cluster.

▶ **Definition 1.1** (Lipschitz decomposition [7]). *Let (X, ρ) be a metric space. A distribution $\mathcal{D}$ over partitions of X is called* (β, Δ)-Lipschitz *if*

1. *for every partition $P \in \operatorname{supp}(\mathcal{D})$, all clusters $C \in P$ satisfy* $\operatorname{diam}(C) \le \Delta$*; and*
2. *for all $x, y \in X$,*

$$\Pr_{P\in\mathcal{D}}[P(x) \neq P(y)] \le \beta \cdot \tfrac{\rho(x,y)}{\Delta},$$

where $P(z)$ denotes the cluster of P containing $z \in X$ and $\operatorname{diam}(C) := \sup_{x,y\in C} \rho(x, y)$.

Typical applications require such decompositions where Δ is not known in advance, or even multiple values of Δ (say for every power of 2). We naturally seek small β and thus define the (optimal) *decomposition parameter of* (X,ρ) as

$$\beta^*(X) := \inf_{\beta\geq 1}\Big\{\beta:\ \forall\Delta>0,\text{ every finite } X'\subseteq X \text{ admits a } (\beta,\Delta)\text{-Lipschitz decomposition}\Big\},$$

and we extend this to a family of metric spaces $\mathcal{X}$, by defining $\beta^*(\mathcal{X}) := \sup_{X\in\mathcal{X}}\beta^*(X)$.

Obtaining bounds on the decomposition parameter of various metrics (and families of metrics) is of significant algorithmic importance, and we list in Table 1 several known bounds. One fundamental example where we know of (nearly) tight bounds is the metric space ℓ_p^d, for $p\geq 1$, which stands for $\mathbb{R}^d$ equipped with the ℓ_p norm. For $p\in[1,2]$, we have $\beta^*(\ell_p^d)=\Theta(d^{1/p})$ due to [13], and for $p\in[2,\infty]$ we have $\beta^*(\ell_p^d)=\tilde{\Theta}(d^{1/2})$ due to [40] (see discussion therein about an incorrect claim made in [13]).[1] Observe that an upper bound for $X=\ell_p^d$ immediately extends to all subsets of it, implying in particular a bound for the family $\mathcal{X}$ of all finite subsets of ℓ_p^d. These bounds depend on d, and are thus most suitable for low-dimensional settings.

We focus on finite metrics X, aiming to bound $\beta^*(X)$ in terms of $n=|X|$, which is often useful in high-dimensional settings. For instance, it is well-known that $\beta^*=\Theta(\log n)$ the family of all n-point metric spaces [7]. To write this assertion more formally, define $\beta_n^*(X) := \beta^*(\{X'\subseteq X:\ |X|=n\})$ and then the above asserts that $\beta_n^*(\ell_\infty)=\Theta(\log n)$, where we used that every finite metric embeds isometrically in ℓ_∞. For the family of n-point ℓ_2 metrics, combining $\beta^*(\ell_2^d)=\tilde{\Theta}(\sqrt{d})$ with the famous JL Lemma [27] immediately yields $\beta_n^*(\ell_2)=O(\sqrt{\log n})$, which is tight by [13]. For n-point ℓ_p metrics, $1<p<2$, we have $\beta_n^*(\ell_p)=\frac{O(\log^{1/p}n)}{p-1}$ due to [35, 40], nearly matching the lower bound of $\beta_n^*(\ell_p)=\Omega(\log^{1/p}n)$ from [13]. However, for n-point ℓ_p metrics, $p>2$, to the best of our knowledge, the only known upper bound is $\beta_n^*(\ell_p)=O(\log n)$, obtained by trivially applying the results for general n-point metric spaces. The following question was raised by Naor [40, Question 1], see also [41, Question 83].

▶ **Question 1.2** ([40]). *Is it true that for every $p\in(2,\infty)$, $\beta_n^*(\ell_p)=o(\log n)$? More ambitiously, is it true that $\beta_n^*(\ell_p)=O_p(\sqrt{\log n})$?*

Our main result, in Theorem 1.3, answers the first part of this question in the affirmative. Additionally, we show in Section 2 an analogous result for another notion of decomposability that was introduced in [22] (and we call capped decomposition) and is particularly suited for high-dimensional geometric spanners.

Geometric Spanners. A *spanner with stretch* $t\geq 1$ (in short a t*-spanner*) for a finite metric $M=(X,\rho)$ is a graph $G=(X,E)$, that satisfies $\rho(x,y)\leq\rho_G(x,y)\leq t\cdot\rho(x,y)$ for all $x,y\in X$, meaning that the shortest-path distance ρ_G in the graph G approximates the original distance $\rho(x,y)$ within factor t, where by definition every edge $\{u,v\}\in E$ has weight $\rho(u,v)$. Of particular interest are spanners that are *sparse*, meaning they contain a small number of edges, ideally linear in $n=|X|$. Another important parameter is the *lightness* of a spanner, defined as the total weight of its edges divided by the weight of a minimum spanning tree of X. Clearly, the lightness is at least 1. These spanners are called geometric because the input is a metric space (rather than a graph). They are natural and useful

[1] Throughout, the notation $\tilde{O}(f)$ hides poly$(\log f)$ factors, and $O_\alpha(\cdot)$ hides a factor that depends only on α.

Table 1 Known bounds on the decomposition parameter of some important families of metrics.

Family of Metrics	β^* or β^*_n	Reference	Comments
ℓ_p^d spaces $1 \le p \le 2$	$\Theta(d^{1/p})$	[13]	
ℓ_p^d spaces $p \ge 2$	$\tilde{\Theta}(\sqrt{d})$	[40]	
finite metrics	$\Theta(\log n)$	[7]	
ℓ_2 space (Euclidean)	$\Theta(\sqrt{\log n})$	[13]	
ℓ_p spaces $1 \le p \le 2$	$\Theta_p(\log^{1/p} n)$	[35, 40]	
ℓ_p spaces $p \ge 2$	$O(\log^{1-1/p} n)$	Theorem 1.3	conjectured $\beta^*_n = \Theta(\sqrt{\log n})$
doubling constant λ	$\Theta(\log \lambda)$	[24]	
K_r-minor-free graphs	$O(r)$	[1, 18]	conjectured $\beta^* = \Theta(\log r)$
graphs with genus g	$\Theta(\log g)$	[36, 1]	
graphs with treewidth w	$\Theta(\log w)$	[20]	

representations of a metric, and as such, have been studied extensively, see the surveys [16, 49, 2]. Spanners for n-point metrics in low-dimensional spaces (e.g., in fixed-dimensional Euclidean space or doubling metrics) are well-studied and well-understood. For instance, metrics with doubling dimension ddim admit $(1+\varepsilon)$-spanners with near-optimal sparsity $n(1/\varepsilon)^{O(\mathrm{ddim})}$ and lightness $(1/\varepsilon)^{O(\mathrm{ddim})}$ [12, 34].

However, in high-dimensional spaces, our understanding of spanners is rather limited. Har-Peled, Indyk, and Sidiropoulos [25] showed that every n-point Euclidean metric admits, an $O(t)$-spanner with $\tilde{O}(n^{1+1/t^2})$ edges for every $t \ge 1$. Filtser and Neiman [22] extended this result to all metric spaces that admit a certain decomposition that we call capped decomposition (Definition 2.4), showing that in those spaces, it is possible to construct spanners that are both sparse and light. In particular, they showed that every n-point subset of ℓ_p, $1 < p \le 2$, has an $O(t)$-spanner with $n^{1+\tilde{O}(1/t^p)}$ edges and lightness $n^{\tilde{O}(1/t^p)}$ for every $t \ge 1$. It remained open whether the spaces ℓ_p for $p \in (2,\infty)$ admit the aforementioned capped decomposition. To the best of our knowledge, all known spanners for these spaces have a tradeoff of stretch $O(t)$ with sparsity $O(n^{1+1/t})$.

1.1 Our Results

Our main contribution is the construction of a Lipschitz decomposition for finite ℓ_p metrics, $p \ge 2$, as follows.

▶ **Theorem 1.3.** *Let $p \in [2,\infty]$. Then $\beta^*_n(\ell_p) = O(\log^{1-1/p} n)$. That is, for every n-point metric $X \subset \ell_p$ and $\Delta > 0$, there exists an $(O(\log^{1-1/p} n), \Delta)$-Lipschitz decomposition of X.*

Previously, this bound was known only for the extreme values $p = 2, \infty$, and in these cases it is actually tight. More precisely, for $p = 2$ our bound coincides with the well-known result $\beta^*_n(\ell_2) = \Theta(\sqrt{\log n})$ [13], and for $p = \Omega(\log n)$ it is known that $\beta^*_n(\ell_p) = \Theta(\log n)$, because all n-point metrics embed into ℓ_p with $O(1)$-distortion [38]. For intermediate values, say fixed $p \in (2,\infty)$, our bound is the first one to improve over $O(\log n)$, which applies to all n-point metrics, and leaves a gap from the $\Omega(\sqrt{\log n})$ lower bound that follows from Dvoretzky's Theorem [15].

We compare our bound with those for other metric spaces in Table 1.

The proof of Theorem 1.3 appears in Section 2.1, and has interesting technical features. It relies on *two* known decompositions of finite metrics, one for general metrics and one for Euclidean metrics, that are composed via a metric-embedding tool called the Mazur map. Our decomposition method is data-dependent, i.e., not oblivious to the data, and we discuss this intriguing aspect in Sections 2.1 and 5.

Note added in proof. Shortly after this work was posted online, two groups working in parallel to each other [31, 42] improved our result in Theorem 1.3 to $\beta_n^*(\ell_p) = O_p(\sqrt{\log n})$ for every $2 < p < \infty$, thereby resolving in the affirmative also the second part of Question 1.2. These two papers design a recursive process that relies on the technique developed here for proving Theorem 1.3, see each paper for its dependence on p.

Geometric Spanners for $p \geq 2$. We then use similar ideas to obtain a new bound for another notion of decomposability, that was introduced in [22] and we call capped decomposition; and this immediately yields geometric spanners in ℓ_p, for $p \geq 2$. While for $p = 2$ these spanners coincide with the known bounds from [25, 22], for fixed $2 < p < \infty$, our spanners are the first improvement over the trivial bounds that hold for all metric spaces.

▶ **Theorem 1.4.** *Let $p \in [2, \infty)$ and $t \geq 1$. Then every n-point metric $X \subset \ell_p$ admits an $O(t)$-spanner of size $\tilde{O}\left(n^{1+1/t^q}\right)$ and lightness $\tilde{O}\left(n^{1/t^q}\right)$, where $q \in (1, 2)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.*

The proof of this theorem appears in Section 2.2, and includes both the spanner construction, which follows [22], and our new bound for capped decomposition, which is the main technical result.

Geometric Spanners for $p \leq 2$. We also sharpen the known spanner results for ℓ_p spaces with $1 < p < 2$, which say that every n-point subset admits an $O(t)$-spanner with $n^{1+O(\log^2 t/t^p)}$ edges and lightness $n^{O(\log^2 t/t^p)}$ for every $t \geq 1$ [22]. We improve upon this result by eliminating the $\log^2 t$ factor in the exponent.

▶ **Theorem 1.5.** *Let $p \in (1, 2]$ and $t \geq 1$. Then every n-point metric $X \subset \ell_p$ admits an $O(t)$-spanner of size $\tilde{O}(n^{1+1/t^p})$ and lightness $\tilde{O}(n^{1/t^p})$.*

The proof of this theorem, presented in Section 3, follows the construction of [22], but replacing a key step, in which they rely on results from [43], with results from [4]. Interestingly, our improved spanner bound "matches" the bounds of Theorem 1.4, up to duality between p and q.

Distance Labeling Schemes. Distance labeling for a metric space (X, ρ) assigns to each point $x \in X$ a label $l(x)$, so that one can later recover (perhaps approximately) the distance between any two points in X based only on their labels (without knowledge of the metric space). It was formulated in [45], motivated by applications in distributed computing, and has been studied intensively, see e.g. [23, 21]. An immediate corollary of our main result in Theorem 1.3 is a distance labeling scheme for finite metrics in ℓ_p for $p > 2$, as follows.

▶ **Theorem 1.6.** *Let $p \in (2, \infty)$. Then the family of n-point metrics in ℓ_p with pairwise distances in the range $[1, \Delta_{max}]$ admits a distance labeling scheme with approximation $O(\log^{1/q} n)$ and label size $O(\log n \log \Delta_{max})$ bits, where $q \in (1, 2)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.*

A formal definition of the distance labeling model and a proof of Theorem 1.6 appear in Section 4.

1.2 Related Work

We focus on Lipschitz decomposition and on capped decomposition, that was introduced in [22], but the literature studies several different decompositions of metric spaces into low-diameter clusters, see e.g. [39, 19]. In particular, the notion of padded decomposition [48, 29] is closely related and was used extensively, see for example [48, 8, 35, 39, 30]. While a Lipschitz decomposition guarantees that nearby points are likely to be clustered together, a padded decomposition guarantees that each point is, with good probability, together with all its nearby points in the same cluster. Remarkably, if a metric space admits a padded decomposition then it admits also a Lipschitz decomposition with almost the same parameters [35], however the other direction is not true, as demonstrated by ℓ_2^d.

The problem of computing efficiently the optimal decomposition parameters for an input metric space (X, ρ) was studied in [32]. Specifically for Lipschitz decomposition, they show that $\beta^*(X)$ can be $O(1)$-approximated in polynomial time (in n).

2 Decompositions and Spanners in ℓ_p for $p > 2$

In this section we consider finite subsets of ℓ_p for $p \in (2, \infty)$. We first present (in Section 2.1) a new Lipschitz decomposition, which proves Theorem 1.3. Next, we show (in Section 2.2) a new construction of capped decomposition, which is a related notion of decomposability that was introduced in [22] without a concrete name. Finally we obtain (in Section 2.3) new spanners, which prove Theorem 1.4. This is actually an immediate corollary of our capped decomposition, by following the spanner construction of [22].

2.1 Lipschitz Decomposition in ℓ_p for $p \in (2, \infty)$

Before presenting the proof of Theorem 1.3, we first provide the intuition behind the proof. A common approach in many algorithms for metric spaces is to embed the given metric into a simpler one (e.g., a tree metric), solve the problem in the target metric, and then pull back this solution to the original metric. For our purpose, of constructing a Lipschitz decomposition of $X \subset \ell_p$, $p > 2$, a natural idea is to seek a low-distortion embedding of X into ℓ_2, because we already have decompositions for that space, namely, $\beta_n^*(\ell_2) = O(\sqrt{\log n})$. Ideally, the embedding into ℓ_2 would be *oblivious*, meaning that it embeds the entire ℓ_p (not only X) into ℓ_2, but unfortunately such an embedding does not exist (it would imply oblivious dimension reduction in ℓ_p for $p > 2$, which is provably impossible [14]). We get around this limitation by employing a *data-dependent* approach, where the decomposition depends on the input set X. More precisely, we use Mazur maps, which provide a low-distortion embedding from ℓ_p to ℓ_2, but only for sets of bounded diameter (see Corollary 2.3). We thus first decompose X into bounded-diameter subsets by applying a standard Lipschitz decomposition (that is applicable for every n-point metric). The final decomposition is obtained by pulling back the solution (clusters) we found in ℓ_2.

We proceed to introduce some technical results needed for our proof of Theorem 1.3. The first one is a well-known bound for Lipschitz decomposition of a finite metric.

▶ **Theorem 2.1** ([7]). *Every n-point metric (X, ρ) admits an $(O(\log n), \Delta)$-Lipschitz decomposition for every $\Delta > 0$.*

Next, we define the *Mazur map*, which is an explicit embedding $M_{p,q} : \ell_p^m \to l_q^m$ for $1 < q < p < \infty$. The image of an input vector v is computed in each coordinate separately, by raising the absolute value to power p/q while keeping the original sign. The next theorem appears in [10], where it is stated as an adaptation of [11], and we will actually need the immediate corollary that follows it.

▶ **Theorem 2.2** ([11, 10]). *Let $1 \le q < p < \infty$ and $C_0 > 0$, and let M be the Mazur map $M_{p,q}$ scaled down by factor $\frac{p}{q}C_0{}^{p/q-1}$. Then for all $x, y \in \ell_p$ such that $||x||_p, ||y||_p \le C_0$,*

$$\tfrac{q}{p}(2C_0)^{1-p/q}||x-y||_p^{p/q} \le ||M(x)-M(y)||_q \le ||x-y||_p.$$

▶ **Corollary 2.3.** *Let $2 < p < \infty$. Every n-point set $X \subset \ell_p$ with diameter at most $C_0 > 0$ admits an embedding $f : X \to \ell_2$ such that*

$$\forall x, y \in X, \qquad \tfrac{2}{p}(2C_0)^{1-p/2}\|x-y\|_p^{p/2} \le \|f(x)-f(y)\|_2 \le \|x-y\|_p.$$

Proof of Theorem 1.3. Let $\Delta > 0$, and let $X \subset \ell_p$ be an n-point metric space for $p \in (2, \infty)$. Construct a partition of X in the following steps:

1. Construct for X an $(O(\log n), \log^{1/p} n \cdot \Delta/4)$-Lipschitz decomposition $P_{\text{init}} = \{K_1, \ldots, K_t\}$ using Theorem 2.1.
2. Embed each cluster $K_i \subset \ell_p$ into ℓ_2 using the embedding f^{K_i} provided by Corollary 2.3 for $C_0 := \log^{1/p} n \cdot \Delta/4$.
3. For each embedded cluster $f^{K_i}(K_i)$, construct an $(O(\sqrt{\log n}), \frac{1}{2}\Delta/\log^{1/2-1/p} n)$-Lipschitz decomposition $P_i = \{K_i^1, \ldots, K_i^{k_i}\}$ using [13] and the JL Lemma [27].
4. The final decomposition P_{out} is obtained by taking the preimage of every cluster of every P_i.

It is easy to see that P_{out} is indeed a partition of X, consisting of $\sum_{i=1}^t k_i$ clusters. Next, consider $x, y \in X$ and let us bound $\Pr[P_{\text{out}}(x) \ne P_{\text{out}}(y)]$. Observe that a pair of points can be separated only in steps 1 or 3. Therefore,

$$\begin{aligned}
&\Pr\Big[P_{\text{out}}(x) \ne P_{\text{out}}(y)\Big] \\
&\le \Pr\Big[P_{\text{init}}(x) \ne P_{\text{init}}(y)\Big] + \Pr\Big[P_i(f^{K_i}(x)) \ne P_i(f^{K_i}(y)) \mid P_{\text{init}}(x) = P_{\text{init}}(y) = K_i\Big] \\
&\le O(\log n)\frac{\|x-y\|_p}{\log^{1/p} n \cdot \Delta/4} + O(\sqrt{\log n})\frac{\|f^{K_i}(x) - f^{K_i}(y)\|_2}{\frac{1}{2}\Delta/\log^{1/2-1/p} n} \\
&\le O(\log^{1-1/p} n)\frac{\|x-y\|_p}{\Delta},
\end{aligned}$$

where the last inequality follows because each f^{K_i} is non-expanding on its cluster $K_i \subset \ell_p$.

It remains to show that the final clusters all have diameter at most Δ. Let $x, y \in X$ be in the same cluster, i.e., $P_{\text{out}}(x) = P_{\text{out}}(y)$. Then $P_{\text{init}}(x) = P_{\text{init}}(y) = K_i$ and $P_i(f^{K_i}(x)) = P_i(f^{K_i}(y))$. Combining the maximum possible diameter of $P_{\text{init}}(x)$ and $P_i(f^{K_i}(x))$ with the contraction guarantees of $f = f^{K_i}$, we get

$$\frac{2}{p}\Big(2(\log^{1/p} n)\frac{\Delta}{4}\Big)^{1-p/2}\|x-y\|_p^{p/2} \le \|f(x)-f(y)\|_2 \le \frac{\Delta}{2}\log^{1/p-1/2} n.$$

Rearranging this, we obtain $\|x-y\|_p \le \frac{\sqrt[p/2]{p/2}}{2}\Delta \le \Delta$, which completes the proof. ◀

2.2 Capped Decomposition in ℓ_p for $p \in (2, \infty)$

We now present our construction of capped decomposition, which is a notion of decomposability that was introduced in [22] without a concrete name. We start with its definition, and then present our construction.

▶ **Definition 2.4.** *Let (X, ρ) be a metric space. A distribution $\mathcal{D}$ over partitions of X is called* (t, Δ, η)-capped *if*

1. *for every partition $P \in \text{supp}(\mathcal{D})$, all clusters $C \in P$ have $\text{diam}(C) \leq \Delta$; and*
2. *for every $x, y \in X$ such that $\rho(x, y) \leq \frac{\Delta}{t}$,*

$$\Pr_{P \in \mathcal{D}}[P(x) = P(y)] \geq \eta.$$

Observe that here, unlike in Lipschitz decomposition, we have a guarantee on the probability that points $x, y \in X$ are clustered together only if they are within distance $\frac{\Delta}{t}$ of each other, hence the name "capped decomposition". Moreover, the probability bound does not depend on the exact value of $\rho(x, y)$. We say that (X, ρ) admits a (t, η)-capped decomposition, where $\eta = \eta(|X|, t)$, if it admits a (t, Δ, η)-capped decomposition for every $\Delta > 0$. A family of metrics admits a (t, η)-capped decomposition if every metric in the family admits a (t, η)-capped decomposition.

▶ **Theorem 2.5.** *Let $p \in (2, \infty)$. Then every n-point metric in ℓ_p admits a $(t, n^{-O(1/t^q)})$-capped decomposition for all $t \geq 1$, where $q \in (1, 2)$ is such that $\frac{1}{p} + \frac{1}{q} = 1$.*

Previously, such a decomposition was known only for the extreme case $p = 2$ by [22], see Proposition 2.6, and our bound above in fact converges to their bound when $p \to 2$. Our proof of Theorem 2.5 is similar to Theorem 1.3, and relies on two known capped decompositions, that we introduce next, together with the Mazur map Corollary 2.3.

▶ **Proposition 2.6** ([22])**.** *Every n-point subset of ℓ_2 admits a $(t, n^{-O(1/t^2)})$-capped decomposition for all $t \geq 1$.*

▶ **Proposition 2.7** (Implicit in [39])**.** *Every n-point metric space admits a $(t, n^{-O(1/t)})$-capped decomposition for all $t \geq 1$.*

Proof of Theorem 2.5. Let $\Delta > 0$ and $t \geq 1$. Let $X \subset \ell_p$ be an n-point subset of $p \in (2, \infty)$, where q is such that $\frac{1}{p} + \frac{1}{q} = 1$. Construct a partition of X in the following steps:

1. Construct for X a $(t_1 := t^q/4, \Delta_1 := \Delta/4t^{1-q}, n^{-O(1/t^q)})$-capped decomposition $P_{\text{init}} = \{K_1, \ldots, K_t\}$ using Proposition 2.7.
2. Embed each cluster $K_i \subset \ell_p$ into ℓ_2 using the embedding f^{K_i} provided by Corollary 2.3 for $C_0 := \Delta_1$.
3. For each embedded cluster $f^{K_i}(K_i)$ construct a $(t_2 := t^{q/2}/2, \Delta_2 := \Delta/2t^{1-q/2}, n^{-O(1/t^q)})$-capped decomposition $P_i = \{K_i^1, \ldots, K_i^{k_i}\}$ using Proposition 2.6.
4. The final decomposition P_{out} is obtained by taking the preimage of every cluster of every P_i.

It is easy to see that that P_{out} is indeed a partition of X, consisting of $\sum_{i=1}^{t} k_i$ clusters. Next, consider $x, y \in X$ with $\|x - y\|_p \leq \Delta/t$ and let us bound $\Pr[P_{\text{out}}(x) = P_{\text{out}}(y)]$. Observe that $\Delta_1/t_1 = \Delta_2/t_2 = \Delta/t$, and therefore

$$\begin{aligned}
&\Pr\Big[P_{\text{out}}(x) = P_{\text{out}}(y)\Big] \\
&\quad = \Pr\Big[P_{\text{init}}(x) = P_{\text{init}}(y)\Big] \cdot \Pr\Big[P_i(f^{K_i}(x)) = P_i(f^{K_i}(y)) \mid P_{\text{init}}(x) = P_{\text{init}}(y) = K_i\Big] \\
&\quad \geq n^{-O(1/t^q)} \cdot n^{-O(1/t^q)} = n^{-O(1/t^q)},
\end{aligned}$$

where the inequality follows because each f^{K_i} is non-expanding on its cluster $K_i \subset \ell_p$.

It remains to show that each cluster has diameter at most Δ. Let $x, y \in X$ be in the same cluster, i.e., $P_{\text{out}}(x) = P_{\text{out}}(y)$. Then $P_{\text{init}}(x) = P_{\text{init}}(y) = K_i$ and $P_i(f^{K_i}(x)) = P_i(f^{K_i}(y))$. Combining the maximum possible diameter of $P_{\text{init}}(x)$ and $P_i(f^{K_i}(x))$ with the contraction guarantees of $f = f^{K_i}$, we get

$$\frac{2}{p}\Big(2\frac{\Delta}{4t^{1-q}}\Big)^{1-p/2}\|x-y\|_p^{p/2} \le \|f(x)-f(y)\|_2 \le \frac{\Delta}{2t^{1-q/2}}.$$

Rearranging this, we obtain $\|x-y\|_p \le \frac{\sqrt[p/2]{p/2}}{2}\Delta \le \Delta$, which completes the proof. ◀

2.3 Spanners in ℓ_p for $p \in (2, \infty)$

We can now prove Theorem 1.4, by applying the following spanner construction of [22].

▶ **Theorem 2.8** ([22]). *Let (X, ρ) be an n-point metric space admitting a (t, η)-capped decomposition for some $t \ge 1$. Then, for every $\epsilon \in (0, 1/8)$, there exists a $(2+\epsilon)t$-spanner for X with $O_\epsilon(\frac{n}{\eta} \cdot \log n \cdot \log t)$ edges and lightness $O_\epsilon(\frac{t}{\eta} \cdot \log^2 n)$.*

Proof of Theorem 1.4. The proof follows directly by combining Theorem 2.5 and Theorem 2.8, as we can assume $t = O(\log n)$ without loss of generality. ◀

3 Spanners in ℓ_p for $p \in (1, 2)$

This section presents an improved construction of geometric spanners in ℓ_p for $p \in (1, 2)$. Previously, $O(t)$-spanners of size $O(n^{1+\log^2 t/t^p})$ for all $t \ge 1$ were constructed in [22]; in particular, setting $t = (\log n \log\log n)^{1/p}$ yields an $O(t)$-spanner of near-linear size $\tilde{O}(n)$. We first present in Section 3.1 two different constructions of near-linear-size spanners with a slightly better stretch. Then in Section 3.2 we use yet another technique, namely Locality Sensitive Hashing (LSH), to slightly improve the construction of [22] of spanners with general stretch $O(t)$.

3.1 Spanners of Near-Linear Size

We slightly improve the near-linear size spanner construction of [22] by shaving the $(\log\log n)^{1/p}$ factor from the stretch, as follows.

▶ **Theorem 3.1.** *For every fixed $p \in (1, 2)$, every n-point metric $X \subset \ell_p$ admits an $O(\log^{1/p} n)$-spanner of size $\tilde{O}(n)$.*

We present two related but different proofs for this theorem. Both are based on modifying the spanner algorithm for ℓ_2 from [25], and therefore we start with an overview of that algorithm. Given an input set $X \subseteq \ell_2$, the algorithm begins by constructing a hierarchical set of 2^i-nets $X = N_0 \supseteq N_1 \supseteq \cdots \supseteq N_{\log \Delta_X}$, where we assume that the minimum and maximum distances in X are 1 and Δ_X, respectively. Then, for each level i, it constructs an $(O(\sqrt{\log n}), O(\sqrt{\log n}) \cdot 2^{i+1})$-Lipschitz decomposition of N_i by combining the JL Lemma [27] with the Lipschitz decomposition of [13]. For each cluster in it, the algorithm add to the spanner edges in a star-like fashion, meaning that all cluster points are connected to one arbitrary point within the cluster. The last two steps are repeated $O(\log n)$ times to ensure that with high probability, for each level i, every $x, y \in N_i$ with $\|x-y\|_2 \le 2^{i+1}$ are clustered together in at least one of the $O(\log n)$ repetitions. It is shown in [25] that this algorithm constructs an $(O(\sqrt{\log n}))$-spanner of size $\tilde{O}(n)$.

Proof of Theorem 3.1 via Lipschitz Decomposition. Observe that the above algorithm of [25] uses the fact that the points lie in ℓ_2 only for the construction of Lipschitz Decompositions, and relies on an optimal decomposition for finite ℓ_2 metrics to conclude that the spanner's stretch is $O(\beta_n^*(\ell_2))$. For finite ℓ_p metrics, $p \in (1,2)$, we can use instead a Lipschitz decomposition from [35, 40], which has $\beta = \frac{O(\log^{1/p} n)}{p-1}$, to conclude the claimed stretch. ◀

We next present a proof that modifies the algorithm of [25] differently, and relies on a decomposition that is similar to a Lipschitz decomposition but has slightly weaker guarantees. Interestingly, this technique yields a slightly stronger result than Theorem 3.1, where p need not to be fixed and can depend on n (e.g., $p \to 1$). We proceed to introduce some technical results from [9] regarding a weak form of dimensionality reduction in ℓ_p, for $p \in [1,2]$, which are needed for our proof.

▶ **Definition 3.2** ([44]). *Let (X,ρ), (Y,τ) be metric spaces and $[a,b]$ be a real interval. An embedding $f : X \to Y$ is called* $[a,b]$-range preserving with distortion $D \geq 1$ *if there exists $c > 0$ such that for all $x, x' \in X$:*

1. *If $a \leq \rho(x,x') \leq b$, then $\rho(x,x') \leq c \cdot \tau(f(x), f(x')) \leq D \cdot \rho(x,x')$.*
2. *If $\rho(x,x') > b$, then $c \cdot \tau(f(x), f(x')) \geq b$.*
3. *If $\rho(x,x') < a$, then $c \cdot \tau(f(x), f(x')) \leq D \cdot a$.*

We say that (X,ρ) admits *an R-range preserving embedding* into (Y,τ) with distortion D, if for all $u > 0$, there exists a $[u, uR]$-range preserving embedding into Y with distortion D.

▶ **Theorem 3.3** ([9]). *Let $1 \leq p \leq 2$. For every n-point set $S \subset \ell_p$, and for every range parameter $R > 1$, there exists an R-range preserving embedding $f : S \to \ell_p^k$ with distortion $1+\epsilon$, such that $k = O\left(\frac{R^{O(1/\epsilon)} \cdot \log n}{\epsilon}\right)$.*

Proof of Theorem 3.1 via Weak Dimension Reduction. Observe that the above algorithm of [25] only requires the decomposition of each net N_i to ensure that points $x, y \in N_i$ with $\|x-y\|_2 \leq 2^{i+1}$ are clustered together with constant probability, and that the diameter of all clusters is at most $O(\sqrt{\log n}) \cdot 2^i$; of course, for $X \subset \ell_p$, $p \in (1,2)$, we replace the $O(\sqrt{\log n})$ factor with $O(\log^{1/p} n)$. A careful examination shows that these properties are preserved by first reducing the dimension using the range-preserving embedding provided by Theorem 3.3 with $\varepsilon = \frac{1}{2}$ and $R = 2$, and then constructing a Lipschitz decomposition for the image points in $\ell_p^{O(\log n)}$ using [13]. ◀

3.2 Spanners with Stretch-Size Tradeoff

We now present, in Theorem 1.5, a construction of $O(t)$-spanners in ℓ_p, where $p \in (1,2)$, of size $\tilde{O}(n^{1+1/t^p})$ for all $t \geq 1$, which slightly improves over the $O(t)$-spanners of size $\tilde{O}(n^{1+\log^2 t/t^p})$ from [22]. It is worth noting that Theorem 1.5 generalizes the results of Theorem 3.1, and thus provides an alternative proof for it.

▶ **Definition 3.4** (LSH [26]). *Let $\mathcal{H}$ be a family of hash functions mapping a metric (X,ρ) to some universe U. We say that $\mathcal{H}$ is* (r, tr, p_1, p_2)-sensitive *if for every $x, y \in X$, the following is satisfied:*

1. *If $\rho(x,y) \leq r$, then $\Pr_{h\in\mathcal{H}}[h(x) = h(y)] \geq p_1$.*
2. *If $\rho(x,y) > tr$, then $\Pr_{h\in\mathcal{H}}[h(x) = h(y)] \leq p_2$.*

Such $\mathcal{H}$ is called an LSH family *with parameter $\gamma := \frac{\log(1/p_1)}{\log(1/p_2)}$.*

▶ **Lemma 3.5** ([22]). *Let (X,ρ) be a metric space such that for every $r > 0$, there exists a (r, tr, p_1, p_2)-sensitive LSH family with parameter γ. Then (X,ρ) admits a $(t, n^{-\mathcal{O}(\gamma)})$-capped decomposition.*

For $p = 2$, the LSH family constructed in [3] can be used in Lemma 3.5 to conclude that ℓ_2 admits a $(t, n^{-O(1/t^2)})$-capped decomposition for every $t \geq 1$ [22], thereby proving Theorem 1.5 for this case of $p = 2$. In a similar fashion, an LSH family constructed in [43] for $p \in (1, 2)$ was used in [22] to show that these spaces admit a $(t, n^{-O(\log^2 t/t^p)})$-capped decomposition. We observe that this result can be improved by replacing the LSH family from [43], with an alternative one that is briefly mentioned in [4], and consequently prove Theorem 1.5. For completeness, we reproduce this LSH family for ℓ_p, where $p \in (1, 2)$.

▶ **Lemma 3.6** ([4]). *Let $p \in (1, 2)$, $r > 0$, and large enough $t > 1$. Then there exists a (r, tr, p_1, p_2)-sensitive LSH family for ℓ_p with parameter $\gamma = \frac{1}{t^p} + o(1)$.*

Proof. Let $p \in (1, 2)$, $r > 0$, and sufficiently large $t > 1$. Let $f : \ell_p \to \ell_2$ be the isometric embedding of the $(p/2)$-snowflake of ℓ_p into ℓ_2 from [28, Theorem 4.1]. Take $r' = r^{p/2}$ and $t' = t^{p/2}$, and let $\mathcal{H}$ be the $(r', t'r', p_1, p_2)$-sensitive LSH family for ℓ_2 with parameter $\gamma = \frac{1}{t'^2} + o(1)$ from [3]. Observe that, for every $x, y \in \ell_p$, if $\|x - y\|_p \leq r$, then $\|f(x) - f(y)\|_2 = \|x - y\|_p^{p/2} \leq r^{p/2} = r'$, and thus

$$\Pr_{h\in\mathcal{H}}[h(f(x)) = h(f(y))] \geq p_1.$$

Similarly, if $\|x - y\|_p > tr$, then $\|f(x) - f(y)\|_2 = \|x - y\|_p^{p/2} > (tr)^{p/2} = t'r'$, and hence

$$\Pr_{h\in\mathcal{H}}[h(f(x)) = h(f(y))] \leq p_2.$$

We therefore conclude that $\mathcal{H} \circ f$ is an (r, tr, p_1, p_2)-sensitive LSH family for ℓ_p with parameter $\gamma = \frac{1}{t^p} + o(1)$. ◀

Proof of Theorem 1.5. The proof follows immediately by constructing a capped decomposition based on Lemma 3.5 and Lemma 3.6, and using it in the spanner construction from Theorem 2.8. ◀

▶ Remark 3.7. While [28, Theorem 4.1] does not provide an efficiently computable embedding, one can compute such an embedding for a finite set of points in polynomial time by [37].

4 Distance Labeling

In the distance labeling model, a scheme is designed for an entire a family $\mathcal{X}$ of n-point metrics (and in some scenarios, all these metrics have the same point set X, e.g., different graphs on the same vertex set). A *scheme* is an algorithm that preprocesses each metric X in $\mathcal{X}$ and assigns to each point $x \in X$ a label $l(x)$.

▶ **Definition 4.1.** *A scheme is* a distance labeling *with approximation $D \geq 1$ and label size of k if*

1. *every label (for every point in every metric in $\mathcal{X}$) consists of at most k bits; and*
2. *there is an algorithm $\mathcal{A}$ that, given the labels $l(x), l(y)$ of two points x, y in a metric $(X, \rho) \in \mathcal{X}$ (but not given (X, ρ) or the points x, y), outputs an estimate $\mathcal{A}(l(x), l(y))$ that satisfies*

$$\rho(x, y) \leq \mathcal{A}(l(x), l(y)) \leq D \cdot \rho(x, y).$$

The following theorem was presented in [24] with limited details, and we include a proof of it below for completeness.

▶ **Theorem 4.2** ([24]). *Let $\mathcal{X}$ be a family of n-point metrics, and assume that all the pairwise distances in all metrics (X, ρ) in $\mathcal{X}$ are in the range $[1, \Delta_{max}]$. Then $\mathcal{X}$ admits a distance-labeling scheme with approximation $O(\beta^*(\mathcal{X}))$ and label size $O(\log n \log \Delta_{\max})$ bits.*

It is straightforward to see that Theorem 1.6 follows by combining Theorem 4.2 and Theorem 1.3.

Proof of Theorem 4.2. We first describe the preprocessing algorithm, denoting $\beta := \beta^*(\mathcal{X})$. Perform the following steps for all levels $i = 0, \ldots, \log \Delta_{\max}$. Begin by constructing a $(\beta, \Delta_i := 4\beta 2^i)$-Lipschitz decomposition, and observe that every two points $x, y \in X$ with $\rho(x, y) \leq 2^i$ are separated with probability at most $\frac{1}{4}$. Then, assign a random bit to each cluster, and observe that if two points are at distance greater than Δ_i, they always fall in different clusters, hence, the probability that they are assigned the same bit is exactly $\frac{1}{2}$, and if they are at distance at most $2^i = \Delta_i/(4\beta)$ they are assigned the same bit with probability at least $\frac{3}{4}$. Repeat the last two steps $k = O(\log n)$ times, and then with high probability, every two points x, y are assigned the same bit at least $\frac{5}{8}k$ times if $\rho(x, y) \leq \Delta_i/(4\beta)$ and fewer than $\frac{5}{8}k$ times if $\rho(x, y) > \Delta_i$. Finally, label each point by concatenating the bit assigned to its cluster in all the repetitions at all levels.

The label-size analysis is straightforward. It remains to show that, given two labels $l(x), l(y)$, it is possible to approximate the distance $\rho(x, y)$ within factor $O(\beta)$. This can be achieved by identifying the smallest level i such that x and y are assigned the same bit at least $\frac{5}{8}k$ times, and then the above analysis (used in contrapositive form) implies that $\Delta_{i-1}/(4\beta) < \rho(x, y) \leq \Delta_i$, where by convention $\Delta_{-1} := 1$. ◀

5 Future Directions

Lipschitz Decompositions. We stress that our decomposition in Theorem 1.3 employs a data-dependent approach, and is not oblivious to the input set X (as, say, the decomposition for ℓ_2 in [13], even when applied together with the JL Lemma). In retrospect, this feature is perhaps not very surprising, because data-dependent approaches have been already shown to be effective for central problems, such as nearest neighbor search [5, 33]. We thus mention that a major open problem in the field is whether dimension reduction is possible in ℓ_p for $p \neq 1, 2, \infty$; we know that for $p > 2$ this is not possible via an oblivious mapping [14], raising the question whether data-dependent mappings can overcome this limitation.

Geometric Spanners. The geometric spanners in [25, 22] for ℓ_p, $1 < p \leq 2$, are not known to be optimal, i.e., we do not know of matching lower bounds, except for the more restricted case of 2-hop spanners [25]. We conjecture that tight instances exist in these spaces, i.e., the spanner bounds obtained in [25, 22] are optimal for every stretch t. We similarly do not know of matching lower bounds for the geometric spanners in ℓ_p, for fixed $2 \leq p < \infty$, that we obtain in Theorem 1.4, and it is quite plausible that our upper bounds are not tight. We do know however, based on known results, that for every n, there exist tight instances in ℓ_p for $p = \Omega(\log n)$.

References

1 Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: padded decomposition for minor-free graphs. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 79–88. Association for Computing Machinery, 2014. doi:10.1145/2591796.2591849.

2 Reyan Ahmed, Greg Bodwin, Faryad Darabi Sahneh, Keaton Hamm, Mohammad Javad Latifi Jebelli, Stephen Kobourov, and Richard Spence. Graph spanners: A tutorial review. *Computer Science Review*, 37:100253, 2020. doi:10.1016/j.cosrev.2020.100253.

3 A. Andoni and P. Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In *47th Annual IEEE Symposium on Foundations of Computer Science*, pages 459–468. IEEE, 2006. doi:10.1109/FOCS.2006.49.

4 A. Andoni and P. Indyk. Nearest neighbors in high-dimensional spaces. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 43, pages 1135–1150. CRC Press, 3rd edition, 2017. doi:10.1201/9781315119601.

5 Alexandr Andoni, Assaf Naor, Aleksandar Nikolov, Ilya P. Razenshteyn, and Erik Waingarten. Data-dependent hashing via nonlinear spectral gaps. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018*, pages 787–800. ACM, 2018. doi:10.1145/3188745.3188846.

6 S. Arora, J. R. Lee, and A. Naor. Euclidean distortion and the sparsest cut. *J. Amer. Math. Soc.*, 21(1):1–21, 2008.

7 Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *37th Annual Symposium on Foundations of Computer Science*, pages 184–193. IEEE, 1996.

8 Y. Bartal. Graph decomposition lemmas and their role in metric embedding methods. In *12th Annual European Symposium on Algorithms*, volume 3221 of *LNCS*, pages 89–97. Springer, 2004. doi:10.1007/978-3-540-30140-0_10.

9 Yair Bartal and Lee-Ad Gottlieb. Dimension reduction techniques for ℓ_p $(1 < p < 2)$ with applications. In *32nd International Symposium on Computational Geometry, SoCG 2016*, volume 51 of *LIPIcs*, pages 16:1–16:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2016. doi:10.4230/LIPICS.SOCG.2016.16.

10 Yair Bartal and Lee-Ad Gottlieb. Approximate nearest neighbor search for ℓ_p-spaces $(2 < p < \infty)$ via embeddings. *Theoretical Computer Science*, 757:27–35, 2019. doi:10.1016/j.tcs.2018.07.011.

11 Yoav Benyamini and Joram Lindenstrauss. *Geometric nonlinear functional analysis*, volume 48. American Mathematical Soc., 1998.

12 Glencora Borradaile, Hung Le, and Christian Wulff-Nilsen. Greedy spanners are optimal in doubling metrics. In *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2371–2379, 2019. doi:10.1137/1.9781611975482.145.

13 M. Charikar, C. Chekuri, A. Goel, S. Guha, and S. Plotkin. Approximating a finite metric by a small number of tree metrics. In *39th Annual Symposium on Foundations of Computer Science*, pages 379–388, 1998.

14 Moses Charikar and Amit Sahai. Dimension reduction in the ℓ_1 norm. In *43rd Symposium on Foundations of Computer Science (FOCS 2002)*, pages 551–560. IEEE Computer Society, 2002. doi:10.1109/SFCS.2002.1181979.

15 A. Dvoretzky. Some results on convex bodies and Banach spaces. In *Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960)*, pages 123–160. Jerusalem Academic Press, Jerusalem, 1961.

16 David Eppstein. Spanning trees and spanners. In *Handbook of Computational Geometry*, pages 425–461. North Holland / Elsevier, 2000. doi:10.1016/B978-044482537-7/50010-3.

17 J. Fakcharoenphol, S. Rao, and K. Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. Syst. Sci.*, 69(3):485–497, 2004. doi:10.1016/j.jcss.2004.04.011.

18 Arnold Filtser. On strong diameter padded decompositions. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019*, volume 145 of *LIPIcs*, pages 6:1–6:21. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICS.APPROX-RANDOM.2019.6.

19 Arnold Filtser. Scattering and sparse partitions, and their applications. *ACM Trans. Algorithms*, 20(4):30:1–30:42, 2024. doi:10.1145/3672562.

20 Arnold Filtser, Tobias Friedrich, Davis Issac, Nikhil Kumar, Hung Le, Nadym Mallek, and Ziena Zeif. Optimal padded decomposition for bounded treewidth graphs. *CoRR*, abs/2407.12230, 2024. `doi:10.48550/arXiv.2407.12230`.

21 Arnold Filtser, Lee-Ad Gottlieb, and Robert Krauthgamer. Labelings vs. embeddings: On distributed and prioritized representations of distances. *Discrete & Computational Geometry*, 71:849–871, 2024. `doi:10.1007/s00454-023-00565-2`.

22 Arnold Filtser and Ofer Neiman. Light spanners for high dimensional norms via stochastic decompositions. *Algorithmica*, 84(10):2987–3007, 2022. `doi:10.1007/s00453-022-00994-0`.

23 C. Gavoille, D. Peleg, S. Pérennes, and R. Raz. Distance labeling in graphs. *Journal of Algorithms*, 53(1):85–112, 2004. `doi:10.1016/J.JALGOR.2004.05.002`.

24 A. Gupta, R. Krauthgamer, and J. R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *44th Annual IEEE Symposium on Foundations of Computer Science*, pages 534–543, 2003. `doi:10.1109/SFCS.2003.1238226`.

25 Sariel Har-Peled, Piotr Indyk, and Anastasios Sidiropoulos. Euclidean spanners in high dimensions. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013*, pages 804–809. SIAM, 2013. `doi:10.1137/1.9781611973105.57`.

26 P. Indyk and R. Motwani. Approximate nearest neighbors: towards removing the curse of dimensionality. In *30th Annual ACM Symposium on Theory of Computing*, pages 604–613, 1998. `doi:10.1145/276698.276876`.

27 W. B. Johnson and J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

28 Nigel J. Kalton. The nonlinear geometry of Banach spaces. *Revista Matematica Complutense*, 21(1):7–60, 2008. URL: `http://eudml.org/doc/42299`.

29 R. Krauthgamer and J. R. Lee. The intrinsic dimensionality of graphs. *Combinatorica*, 27(5):551–585, 2007. `doi:10.1007/s00493-007-2183-y`.

30 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. `doi:10.1007/s00039-005-0527-6`.

31 Robert Krauthgamer, Nir Petruschka, and Shay Sapir. The power of recursive embeddings for ℓ_p metrics, 2025. `doi:10.48550/arXiv.2503.18508`.

32 Robert Krauthgamer and Tim Roughgarden. Metric clustering via consistent labeling. *Theory of Computing*, 7(5):49–74, 2011. `doi:10.4086/toc.2011.v007a005`.

33 Deepanshu Kush, Aleksandar Nikolov, and Haohua Tang. Near neighbor search via efficient average distortion embeddings. In *37th International Symposium on Computational Geometry, SoCG 2021*, volume 189 of *LIPIcs*, pages 50:1–50:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021. `doi:10.4230/LIPICS.SOCG.2021.50`.

34 Hung Le and Shay Solomon. A unified framework for light spanners. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023*, pages 295–308. ACM, 2023. `doi:10.1145/3564246.3585185`.

35 J. R. Lee and A. Naor. Metric decomposition, smooth measures, and clustering. Unpublished manuscript, 2003. URL: `https://web.math.princeton.edu/~naor/homepagefiles/cluster.pdf`.

36 James R. Lee and Anastasios Sidiropoulos. Genus and the geometry of the cut graph. In *21st Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 193–201. SIAM, 2010. `doi:10.1137/1.9781611973075.18`.

37 N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995. `doi:10.1007/BF01200757`.

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

39 M. Mendel and A. Naor. Ramsey partitions and proximity data structures. *J. Eur. Math. Soc.*, 9(2):253–275, 2007.

40 Assaf Naor. Probabilistic clustering of high dimensional norms. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017*, pages 690–709. SIAM, 2017. doi:10.1137/1.9781611974782.44.

41 Assaf Naor. *Extension, separation and isomorphic reverse isoperimetry*, volume 11 of *Mem. Eur. Math. Soc.* European Mathematical Society (EMS), 2024. doi:10.4171/MEMS/11.

42 Assaf Naor and Kevin Ren. ℓ_p has nontrivial Euclidean distortion growth when $2 < p < 4$, 2025. arXiv:2502.10543.

43 Huy L. Nguyen. Approximate nearest neighbor search in ℓ_p. *CoRR*, abs/1306.3601, 2013. arXiv:1306.3601.

44 R. Ostrovsky and Y. Rabani. Polynomial-time approximation schemes for geometric min-sum median clustering. *J. ACM*, 49(2):139–156, 2002. doi:10.1145/506147.506149.

45 D. Peleg. Proximity-preserving labeling schemes. *Journal of Graph Theory*, 33(3):167–176, 2000. doi:10.1002/(SICI)1097-0118(200003)33:3\%3C167::AID-JGT7\%3E3.0.CO;2-5.

46 D. Peleg and A. A. Schäffer. Graph spanners. *J. Graph Theory*, 13(1):99–116, 1989. doi:10.1002/jgt.3190130114.

47 David Peleg and Jeffrey D. Ullman. An optimal synchronizer for the hypercube. *SIAM J. Comput.*, 18(4):740–747, 1989. doi:10.1137/0218050.

48 S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the 15th Annual Symposium on Computational Geometry*, pages 300–306. ACM, 1999. doi:10.1145/304893.304983.

49 Uri Zwick. Exact and approximate distances in graphs – A survey. In *Algorithms – ESA 2001, 9th Annual European Symposium*, volume 2161 of *Lecture Notes in Computer Science*, pages 33–48. Springer, 2001. doi:10.1007/3-540-44676-1_3.