Light Spanners for High Dimensional Norms via Stochastic Decompositions

Arnold Filtser[1]
Ben-Gurion University of the Negev ,Beer-Sheva, Israel
arnoldf@cs.bgu.ac.il

Ofer Neiman[2]
Ben-Gurion University of the Negev, Beer-Sheva, Israel
neimano@cs.bgu.ac.il

Abstract

Spanners for low dimensional spaces (e.g. Euclidean space of constant dimension, or doubling metrics) are well understood. This lies in contrast to the situation in high dimensional spaces, where except for the work of Har-Peled, Indyk and Sidiropoulos (SODA 2013), who showed that any n-point Euclidean metric has an $O(t)$-spanner with $\tilde{O}(n^{1+1/t^2})$ edges, little is known.

In this paper we study several aspects of spanners in high dimensional normed spaces. First, we build spanners for finite subsets of ℓ_p with $1 < p \le 2$. Second, our construction yields a spanner which is both sparse and also *light*, i.e., its total weight is not much larger than that of the minimum spanning tree. In particular, we show that any n-point subset of ℓ_p for $1 < p \le 2$ has an $O(t)$-spanner with $n^{1+\tilde{O}(1/t^p)}$ edges and lightness $n^{\tilde{O}(1/t^p)}$.

In fact, our results are more general, and they apply to any metric space admitting a certain low diameter stochastic decomposition. It is known that arbitrary metric spaces have an $O(t)$-spanner with lightness $O(n^{1/t})$. We exhibit the following tradeoff: metrics with decomposability parameter $\nu = \nu(t)$ admit an $O(t)$-spanner with lightness $\tilde{O}(\nu^{1/t})$. For example, n-point Euclidean metrics have $\nu \le n^{1/t}$, metrics with doubling constant λ have $\nu \le \lambda$, and graphs of genus g have $\nu \le g$. While these families do admit a $(1+\epsilon)$-spanner, its lightness depend exponentially on the dimension (resp. $\log g$). Our construction alleviates this exponential dependency, at the cost of incurring larger stretch.

2012 ACM Subject Classification CCS → Theory of computation → Design and analysis of algorithms → Graph algorithms analysis → Sparsification and spanners

Keywords and phrases Spanners, Stochastic Decompositions, High Dimensional Euclidean Space, Doubling Dimension, Genus Graphs

Digital Object Identifier 10.4230/LIPIcs.ESA.2018.29

1 Introduction

1.1 Spanners

Given a metric space (X, d_X), a weighted graph $H = (X, E)$ is a t-*spanner* of X, if for every pair of points $x, y \in X$, $d_X(x, y) \le d_H(x, y) \le t \cdot d_X(x, y)$ (where d_H is the shortest path metric in H). The factor t is called the *stretch* of the spanner. Two important parameters of interest are: the *sparsity* of the spanner, i.e. the number of edges, and the *lightness* of the

[1] Partially supported by the Lynn and William Frankel Center for Computer Sciences, ISF grant 1817/17, and by BSF Grant 2015813.
[2] Partially supported by ISF grant 1817/17, and by BSF Grant 2015813.

26th Annual European Symposium on Algorithms (ESA 2018).
Editors: Yossi Azar, Hannah Bast, and Grzegorz Herman; Article No. 29; pp. 29:1–29:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

spanner, which is the ratio between the total weight of the spanner and the weight of the minimum spanning tree (MST).

The tradeoff between stretch and sparsity/lightness of spanners is the focus of an intensive research effort, and low stretch spanners were used in a plethora of applications, to name a few: Efficient broadcast protocols [8, 9], network synchronization [6, 49, 8, 9, 48], data gathering and dissemination tasks [14, 60, 22], routing [61, 49, 50, 57], distance oracles and labeling schemes [47, 58, 53], and almost shortest paths [19, 52, 23, 25, 28].

Spanners for general metric spaces are well understood. The seminal paper of [4] showed that for any parameter $k \geq 1$, any metric admits a $(2k-1)$-spanner with $O(n^{1+1/k})$ edges, which is conjectured to be best possible. For light spanners, improving [17, 24], it was shown in [18] that for every constant $\epsilon > 0$ there is a $(2k-1)(1+\epsilon)$-spanner with lightness $O(n^{1/k})$ and at most $O(n^{1+1/k})$ edges.

There is an extensive study of spanners for restricted classes of metric spaces, most notably subsets of low dimensional Euclidean space, and more generally doubling metrics.[3] For such low dimensional metrics, much better spanners can be obtained. Specifically, for n points in d-dimensional Euclidean space, [54, 59, 21] showed that for any $\epsilon \in (0, \frac{1}{2})$ there is a $(1+\epsilon)$-spanner with $n \cdot \epsilon^{-O(d)}$ edges and lightness $\epsilon^{-O(d)}$ (further details on Euclidean spanners could be found in [45]). This result was recently generalized to doubling metrics by [12], with $\epsilon^{-O(\mathrm{ddim})}$ lightness and $n \cdot \epsilon^{-O(\mathrm{ddim})}$ edges (improving [55, 30, 29]). Such low stretch spanners were also devised for metrics arising from certain graph families. For instance, [4] showed that any planar graph admits a $(1+\epsilon)$-spanner with lightness $O(1/\epsilon)$. This was extended to graphs with small genus[4] by [31], who showed that every graph with genus $g > 0$ admits a spanner with stretch $(1+\epsilon)$ and lightness $O(g/\epsilon)$. A long sequence of works for other graph families, concluded recently with a result of [13], who showed $(1+\epsilon)$-spanners for graphs excluding K_r as a minor, with lightness $\approx O(r/\epsilon^3)$.

In all these results there is an exponential dependence on a certain parameter of the input metric space (the dimension, the logarithm of the genus/minor-size), which is unfortunately unavoidable for small stretch (for all n-point metric spaces the dimension/parameter is at most $O(\log n)$, while spanner with stretch better than 3 requires in general $\Omega(n^2)$ edges [58]). So when the relevant parameter is small, light spanners could be constructed with stretch arbitrarily close to 1. However, in metrics arising from actual data, the parameter of interest may be moderately large, and it is not known how to construct light spanners avoiding the exponential dependence on it. In this paper, we devise a tradeoff between stretch and sparsity/lightness that can diminish this exponential dependence. To the best of our knowledge, the only such tradeoff is the recent work of [34], who showed that n-point subsets of Euclidean space (in any dimension) admit a $O(t)$-spanner with $\tilde{O}(n^{1+1/t^2})$ edges (without any bound on the lightness).

1.2 Stochastic Decompositions

In a (stochastic) decomposition of a metric space, the goal is to find a partition of the points into clusters of low diameter, such that the probability of nearby points to fall into different clusters is small. More formally, for a metric space (X, d_X) and parameters $t \geq 1$

[3] A metric space (X, d) has doubling constant λ if for every $x \in X$ and radius $r > 0$, the ball $B(x, 2r)$ can be covered by λ balls of radius r. The doubling dimension is defined as $\mathrm{ddim} = \log_2 \lambda$. A d-dimensional ℓ_p space has $\mathrm{ddim} = \Theta(d)$, and every n point metric has $\mathrm{ddim} = O(\log n)$.

[4] The *genus* of a graph is minimal integer g, such that the graph could be drawn on a surface with g "handles".

and $\delta = \delta(|X|, t) \in [0, 1]$, we say that the metric is (t, δ)-decomposable, if for every $\Delta > 0$ there is a probability distribution over partitions of X into clusters of diameter at most $t \cdot \Delta$, such that every two points of distance at most Δ have probability at least δ to be in the same cluster.

Such decompositions were introduced in the setting of distributed computing [7, 43], and have played a major role in the theory of metric embedding [10, 51, 26, 38, 39, 1], distance oracles and routing [44, 2], multi-commodity flow/sparsest cut gaps [41, 37] and also were used in approximation algorithms and spectral methods [15, 36, 11]. We are not aware of any direct connection of these decompositions to spanners (except spanners for general metrics implicit in [44, 2]).

Note that our definition is slightly different than the standard one. The probability δ that a pair $x, y \in X$ is in the same cluster may depend on $|X|$ and t, but unlike previous definitions, it does not depend on the precise value of $d_X(x, y)$ (rather, only on the fact that it is bounded by Δ). This simplification suits our needs, and it enables us to capture more succinctly the situation for high dimensional normed spaces, where the dependence of δ on $d_X(x, y)$ is non-linear. These stochastic decompositions are somewhat similar to Locality Sensitive Hashing (LSH), that were used by [34] to construct spanners. The main difference is that in LSH, far away points may be mapped to the same cluster with some small probability, and more focus was given to efficient computation of the hash function. It is implicit in [34] that existence of good LSH imply sparse spanners.

A classic tool for constructing spanners in normed and doubling spaces is WSPD (Well Separated Pair Decomposition, see [16, 56, 35]). Given a set of points P, a WSPD is a set of pairs $\{(A_i, B_i)\}_i$ of subsets of P, where the diameters of A_i and B_i are at most an ϵ-fraction of $d(A_i, B_i)$, and such that for every pair $x, y \in P$ there is some i with $(x, y) \in A_i \times B_i$. A WSPD is designed to create a $(1 + O(\epsilon))$-spanner, by adding an arbitrary edge between a point in A_i and a point in B_i for every i (as opposed to our construction, based on stochastic decompositions, in which we added only inner-cluster edges). An exponential dependence on the dimension is unavoidable with such a low stretch, thus it is not clear whether one can use a WSPD to obtain very sparse or light spanners in high dimensions.

1.3 Our Results

Our main result is exhibiting a connection between stochastic decompositions of metric spaces, and light spanners. Specifically, we show that if an n-point metric is (t, δ)-decomposable, then for any constant $\epsilon > 0$, it admits a $(2 + \epsilon) \cdot t$-spanner with $\tilde{O}(n/\delta)$ edges and lightness $\tilde{O}(1/\delta)$. (Abusing notation, $\tilde{O}$ hides $\text{polylog}(n)$ factors.)

It can be shown that Euclidean metrics are $(t, n^{-O(1/t^2)})$-decomposable, thus our results extends [34] by providing a smaller stretch $(2 + \epsilon) \cdot t$-spanner, which is both sparse – with $\tilde{O}(n^{1+O(1/t^2)})$ edges – and has lightness $\tilde{O}(n^{O(1/t^2)})$. For d-dimensional Euclidean space, where $d = o(\log n)$ we can obtain lightness $\tilde{O}(2^{O(d/t^2)})$ and $\tilde{O}(n \cdot 2^{O(d/t^2)})$ edges. We also show that n-point subsets of ℓ_p spaces for any fixed $1 < p < 2$ are $(t, n^{-O(\log^2 t/t^p)})$-decomposable, which yields light spanners for such metrics as well.

In addition, metrics with doubling constant λ are $(t, \lambda^{-O(1/t)})$-decomposable [33, 1], and graphs with genus g are $(t, g^{-O(1/t)})$-decomposable [40, 3], which enables us to alleviate the exponential dependence on ddim and $\log g$ in the sparsity/lightness by increasing the stretch. See Table 1 for more details. (We remark that for graphs excluding K_r as a minor, the current best decomposition achieves probability only $2^{-O(r/t)}$ [3]; if this will be improved to the conjectured $r^{-O(1/t)}$, then our results would provide interesting spanners for this family as well.)

	Stretch	Lightness	Sparsity	
Euclidean space	$O(t)$	$\tilde{O}(n^{1/t^2})$	$\tilde{O}(n^{1+1/t^2})$	Corollary 6
	$O(\sqrt{\log n})$	$\tilde{O}(1)$	$\tilde{O}(n)$	
ℓ_p space, $1 < p < 2$	$O(t)$	$\tilde{O}(n^{\log^2 t/t^p})$	$\tilde{O}(n^{1+\log^2 t/t^p})$	Corollary 7
	$O((\log n \cdot \log\log n)^{1/p})$	$\tilde{O}(1)$	$\tilde{O}(n)$	
Doubling constant λ	$O(t)$	$\tilde{O}(\lambda^{1/t})$	$\tilde{O}(n \cdot \lambda^{1/t})$	Corollary 8
	$O(\log \lambda)$	$\tilde{O}(1)$	$\tilde{O}(n)$	
Graph with genus g	$O(t)$	$\tilde{O}(g^{1/t})$	$O(n+g)$	Corollary 9
	$O(\log g)$	$\tilde{O}(1)$	$O(n+g)$	

Table 1 In this table we summarize some corollaries of our main result. The metric spaces have cardinality n, and $\tilde{O}$ hides (mild) polylog(n) factors. The stretch t is a parameter ranging between 1 and $\log n$.

Note that up to polylog(n) factors, our stretch-lightness tradeoff generalizes the [18] spanner for general metrics, which has stretch $(2t-1)(1+\epsilon)$ and lightness $O(n^{1/t})$. Define for a (t,δ)-decomposable metric the parameter $\nu = 1/\delta^t$. Then we devise for such a metric a $(2t-1)(1+\epsilon)$-spanner with lightness $O(\nu^{1/t})$.

For example, consider an n-point metric with doubling constant $\lambda = 2^{\sqrt{\log n}}$. No spanner with stretch $o(\log n/\log\log n)$ and lightness $\tilde{O}(1)$ for such a metric was known. Our result implies such a spanner, with stretch $O(\sqrt{\log n})$.

We also remark that the existence of light spanners does not imply decomposability. For example, consider the shortest path metrics induced by bounded-degree expander graphs. Even though these metrics have the (asymptotically) worst possible decomposability parameters (they are only $(t, n^{-\Omega(1/t)})$-decomposable [42]), they nevertheless admit 1-spanners with constant lightness (the spanner being the expander graph itself).

2 Preliminaries

Given a metric space (X, d_X), let T denote its minimum spanning tree (MST) of weight L. For a set $A \subseteq X$, the diameter of A is $\mathrm{diam}(A) = \max_{x,y\in A} d_X(x,y)$. Assume, as we may, that the minimal distance in X is 1.

By O_ϵ we denote asymptotic notation which hides polynomial factors of $\frac{1}{\epsilon}$, that is $O_\epsilon(f) = O(f)\cdot \mathrm{poly}(\frac{1}{\epsilon})$. Unless explicitly specified otherwise, all logarithms are in base 2.

Nets. For $r > 0$, a set $N \subseteq X$ is an *r-net*, if (1) for every $x \in X$ there is a point $y \in N$ with $d_X(x,y) \le r$, and (2) every pair of net points $y, z \in N$ satisfy $d_X(y,z) > r$. It is well known that nets can be constructed in a greedy manner. For $0 < r_1 \le r_2 \le \cdots \le r_s$, a *hierarchical net* is a collection of nested sets $X \supseteq N_1 \supseteq N_2 \supseteq \cdots \supseteq N_s$, where each N_i is an r_i-net. Since N_{i+1} satisfies the second condition of a net with respect to radius r_i, one can obtain N_i from N_{i+1} by greedily adding points until the first condition is satisfied as well. In the following claim we argue that nets are sparse sets with respect to the MST weight.

▶ **Claim 1.** Consider a metric space (X, d_X) with MST of weight L, let N be an r-net, then $|N| \le \frac{2L}{r}$.

Proof. Let T be the MST of X, note that for every $x, y \in N$, $d_T(x,y) \ge d_X(x,y) > r$. For a point $x \in N$, $B_T(x,b) = \{y \in X \mid d_T(x,y) \le b\}$ is the ball of radius b around x in the MST metric. We say that an edge $\{y,z\}$ of T is *cut* by the ball $B_T(x,b)$ if $d_T(x,y) < b < d_T(x,z)$.

Consider the set $\mathcal{B}$ of balls of radius $r/2$ around the points of N. We can subdivide[5] the edges of T until no edge is cut by any of the balls of $\mathcal{B}$. Note that the subdivisions do not change the total weight of T nor the distances between the original points of X.

If both the endpoints of an edge e belong to the ball B, we say that the edge e is internal to B. By the second property of nets, and since $B_T(x,b) \subseteq B_X(x,b)$, the set of internal edges corresponding to the balls $\mathcal{B}$ are disjoint. On the other hand, as the tree is connected, the weight of the internal edges in each ball must be at least $r/2$. As the total weight is bounded by L, the claim follows. ◀

Stochastic Decompositions. Consider a *partition* $\mathcal{P}$ of X into disjoint clusters. For $x \in X$, we denote by $\mathcal{P}(x)$ the cluster $P \in \mathcal{P}$ that contains x. A partition $\mathcal{P}$ is Δ-*bounded* if for every $P \in \mathcal{P}$, $\text{diam}(P) \leq \Delta$. If a pair of points x, y belong to the same cluster, i.e. $\mathcal{P}(x) = \mathcal{P}(y)$, we say that they are *clustered* together by $\mathcal{P}$.

▶ **Definition 2.** For metric space (X, d_X) and parameters $t \geq 1$, $\Delta > 0$ and $\delta \in [0,1]$, a distribution $\mathcal{D}$ over partitions of X is called a (t, Δ, δ)-decomposition, if it fulfills the following properties.

- Every $\mathcal{P} \in \text{supp}(\mathcal{D})$ is $t \cdot \Delta$-bounded.
- For every $x, y \in X$ such that $d_X(x,y) \leq \Delta$, $\Pr_{\mathcal{D}}[\mathcal{P}(x) = \mathcal{P}(y)] \geq \delta$.

A metric is (t, δ)-decomposable, where $\delta = \delta(|X|, t)$, if it admits a (t, Δ, δ)-decomposition for any $\Delta > 0$. A family of metrics is (t, δ)-decomposable if each member (X, d_X) in the family is (t, δ)-decomposable.

We observe that if a metric (X, d_X) is $(t, \delta(|X|, t))$-decomposable, then also every sub-metric $Y \subseteq X$ is $(t, \delta(|X|, t))$-decomposable. In some cases Y is also $(t, \delta(|Y|, t))$-decomposable (we will exploit these improved decompositions for subsets of ℓ_p). The following claim argues that sampling $O(\frac{\log n}{\delta})$ partitions suffices to guarantee that every pair is clustered at least once.

▶ **Claim 3.** Let (X, d_X) be a metric space which admits a (t, Δ, δ)-decomposition, and let $N \subseteq X$ be of size $|N| = n$. Then there is a set $\{\mathcal{P}_1, \ldots, \mathcal{P}_\varphi\}$ of $t \cdot \Delta$-bounded partitions of N, where $\varphi = \frac{2 \ln n}{\delta}$, such that every pair $x, y \in N$ at distance at most Δ is clustered together by at least one of the $\mathcal{P}_i$.

Proof. Let $\{\mathcal{P}_1, \ldots, \mathcal{P}_\varphi\}$ be i.i.d partitions drawn from the (t, Δ, δ)-decomposition of X. Consider a pair $x, y \in N$ at distance at most Δ. The probability that x, y are not clustered in any of the partitions is bounded by

$$\Pr[\forall i, \quad \mathcal{P}_i(x) \neq \mathcal{P}_i(y)] \leq (1-\delta)^{(2 \ln n)/\delta} \leq \frac{1}{n^2} \,.$$

The claim now follows by the union bound. ◀

3 Light Spanner Construction

In this section we present a generalized version of the algorithm of [34], depicted in Algorithm 1. The differences in execution and analysis are: (1) Our construction applies to general

[5] To subdivide an edge $e = \{x, y\}$ of weight w the following steps are taken: (1) Delete the edge e. (2) Add a new vertex v_e. (3) Add two new edges $\{x, v_e\}, \{v_e, y\}$ with weights $\alpha \cdot w$ and $(1-\alpha) \cdot w$ for some $\alpha \in (0,1)$.

decomposable metric spaces – we use decompositions rather than LSH schemes. (2) We analyze the lightness of the resulting spanners. (3) We achieve stretch $t \cdot (2+\epsilon)$ rather than $O(t)$.

The basic idea is as follows. For every weight scale $\Delta_i = (1+\epsilon)^i$, construct a sequence of $t \cdot \Delta_i$-bounded partitions $\mathcal{P}_1, \dots, \mathcal{P}_\varphi$ such that every pair x, y at distance $\le \Delta_i$ will be clustered together at least once. Then, for each $j \in [\varphi]$ and every cluster $P \in \mathcal{P}_j$, we pick an arbitrary root vertex $v_P \in P$, and add to our spanner edges from v_P to all the points in P. This ensures stretch $2t \cdot (1+\epsilon)$ for all pairs with $d_X(x,y) \in [(1-\epsilon)\Delta_i, \Delta_i]$. Thus, repeating this procedure on all scales $i = 1, 2, \dots$ provides a spanner with stretch $2t \cdot (1+\epsilon)$.

However, the weight of the spanner described above is unbounded. In order to address this problem at scale Δ_i, instead of taking the partitions over all points, we partition only the points of an $\epsilon\Delta_i$-net. The stretch is still small: x, y at distance Δ_i will have nearby net points $\tilde{x}, \tilde{y}$. Then, a combination of newly added edges with older ones will produce a short path between x to y. The bound on the lightness will follow from the observation that the number of net points is bounded with respect to the MST weight.

▶ **Theorem 4.** *Let (X, d_X) be a (t, δ)-decomposable n-point metric space. Then for every $\epsilon \in (0, 1/8)$, there is a $t\cdot(2+\epsilon)$-spanner for X with lightness $O_\epsilon\left(\frac{t}{\delta} \cdot \log^2 n\right)$ and $O_\epsilon\left(\frac{n}{\delta} \cdot \log n \cdot \log t\right)$ edges.*

Algorithm 1 $H =$ `Spanner-From-Decompositions`$((X, d_X), t, \epsilon)$

```
Let N_0 ⊇ N_1 ⊇ ··· ⊇ N_{log_{1+ε} L} be a hierarchical net, where N_i is ε·Δ_i = ε·(1+ε)^i-net
of (X, d_X).
for i ∈ {0, 1, ..., log_{1+ε} L} do
  For parameters Δ = (1+2ε)Δ_i and t, let P_1, ..., P_{φ_i} be the set of t·Δ-bounded
  partitions guaranteed by Claim 3 on the set N_i.
  for j ∈ {1, ..., φ_i} and P ∈ P_j do
    Let v_P ∈ P be an arbitrarily point.
    Add to H an edge from every point x ∈ P \ {v_P} to v_P.
  end for
end for
return H.
```

Proof. We will prove stretch $t \cdot (2 + O(\epsilon))$ instead of $t \cdot (2+\epsilon)$. This is good enough, as post factum we can scale ϵ accordingly.

Stretch Bound. Let $c > 1$ be a constant (to be determined later). Consider a pair $x, y \in X$ such that $(1+\epsilon)^{i-1} < d_X(x,y) \le (1+\epsilon)^i$. We will assume by induction that every pair x', y' at distance at most $(1+\epsilon)^{i-1}$ already enjoys stretch at most $\alpha = t \cdot (2 + c \cdot \epsilon)$ in H. Set $\Delta_i = (1+\epsilon)^i$, and let $\tilde{x}, \tilde{y} \in N_i$ be net points such that $d_X(x, \tilde{x}), d_X(y, \tilde{y}) \le \epsilon \cdot \Delta_i$. By the triangle inequality $d_X(\tilde{x}, \tilde{y}) \le (1+2\epsilon) \cdot \Delta_i = \Delta$. Therefore there is a $t \cdot \Delta$-bounded partition $\mathcal{P}$ constructed at round i such that $\mathcal{P}(\tilde{x}) = \mathcal{P}(\tilde{y})$. In particular, there is a center vertex $v = v_{\mathcal{P}(\tilde{x})}$ such that both $\{\tilde{x}, v\}, \{\tilde{y}, v\}$ were added to the spanner H. Using the induction hypothesis on the pairs $\{x,\tilde{x}\}$ and $\{y,\tilde{y}\}$, we conclude

$$\begin{aligned} d_H(x,y) &\le d_H(x,\tilde{x}) + d_H(\tilde{x},v) + d_H(v,\tilde{y}) + d_H(\tilde{y},y) \\ &\le \alpha\cdot\epsilon\Delta_i + (1+2\epsilon)t\Delta_i + (1+2\epsilon)t\Delta_i + \alpha\cdot\epsilon\Delta_i \\ &\overset{(*)}{<} \frac{\alpha}{1+\epsilon}\cdot\Delta_i \le \alpha\cdot d_X(x,y)\ , \end{aligned}$$

where the inequality $(*)$ follows as $2(1+2\epsilon)t < \alpha(\frac{1}{1+\epsilon} - 2\epsilon)$ for large enough constant c, using that $\epsilon < 1/8$.

Sparsity bound. For a point $x \in X$, let s_x be the maximal index such that $x \in N_{s_x}$. Note that the number of edges in our spanner is not affected by the choice of "cluster centers" in line 5 in Algorithm 1. Therefore, the edge count will be still valid if we assume that $v_P \in P$ is the vertex y with maximal value s_y among all vertices in P.

Consider an edge $\{x,y\}$ added during the i's phase of the algorithm. Necessarily $x,y \in N_i$, and x,y belong to the same cluster P of a partition $\mathcal{P}_j$. W.l.o.g, $y = v_P$, in particular $s_x \le s_y$. The edge $\{x,y\}$ will be charged upon x. Since the partitions at level i are $t\cdot\Delta$ bounded, we have that $d_X(x,y) \le t\cdot\Delta = t\cdot(1+2\epsilon)\cdot(1+\epsilon)^i$. Hence, for i' such that $\epsilon\cdot(1+\epsilon)^{i'} > t\cdot(1+2\epsilon)\cdot(1+\epsilon)^i$, i.e. $i' > i + O_\epsilon(\log t)$, the points x,y cannot both belong to $N_{i'}$. As $s_x \le s_y$, it must be that $x \notin N_{i'}$. We conclude that x can be charged in at most $O_\epsilon(\log t)$ different levels. As in level i each vertex is charged for at most $\varphi_i \le O(\frac{\log n}{\delta})$ edges, the total charge for each vertex is bounded by $O_\epsilon(\frac{\log n\cdot\log t}{\delta})$.

Lightness bound. Consider the scale $\Delta_i = (1+\epsilon)^i$. As N_i is an $\epsilon\cdot\Delta_i$-net, Claim 1 implies that N_i has size $n_i \le \frac{2L}{\epsilon\cdot\Delta_i}$, and in any case at most n. In that scale, we constructed $\varphi_i = \frac{2}{\delta}\log n_i \le \frac{2}{\delta}\log n$ partitions, adding at most n_i edges per partition. The weight of each edge added in this scale is bounded by $O(t\cdot\Delta_i)$.

Let H_1 consist of all the edges added in scales $i \in \{\log_{1+\epsilon}\frac{L}{n},\dots,\log_{1+\epsilon}L\}$, while H_2 consist of edges added in the lower scales. Note that $H = H_1 \cup H_2$.

$$\begin{aligned} w(H_1) &\le \sum_{i\in\left\{\log_{1+\epsilon}\frac{L}{n},\dots,\log_{1+\epsilon}L\right\}} O(t\cdot\Delta_i)\cdot n_i\cdot\varphi_i \\ &= O\left(\frac{t}{\delta}\cdot\log n\cdot\sum_{i\in\left\{\log_{1+\epsilon}\frac{L}{n},\dots,\log_{1+\epsilon}L\right\}}\Delta_i\cdot\frac{L}{\epsilon\cdot\Delta_i}\right) = O_\epsilon\left(\frac{t}{\delta}\cdot\log^2 n\right)\cdot L\ . \\ w(H_2) &\le \sum_{\Delta_i\in\frac{L}{n}\cdot\{(1+\epsilon)^{-1},(1+\epsilon)^{-2},\dots,\}} O(t\cdot\Delta_i)\cdot n_i\cdot\varphi_i \\ &= O\left(\frac{t}{\delta}\cdot\log n\cdot\sum_{i\ge 1}\frac{1}{(1+\epsilon)^i}\right)\cdot L = O_\epsilon\left(\frac{t}{\delta}\cdot\log n\right)\cdot L\ . \end{aligned}$$

The bound on the lightness follows. ◀

4 Corollaries and Extensions

In this section we describe some corollaries of Theorem 4 for certain metric spaces, and show some extensions, such as improved lightness bound for normed spaces, and discuss graph spanners.

4.1 High Dimensional Normed Spaces

Here we consider the case that the given metric space (X, d) satisfies that every sub-metric $Y \subseteq X$ of size $|Y| = n$ is (t, δ)-decomposable for $\delta = n^{-\beta}$, where $\beta = \beta(t) \in (0, 1)$ is a function of t. In such a case we are able to shave a $\log n$ factor in the lightness.

▶ **Theorem 5.** *Let (X, d_X) be an n-point metric space such that every $Y \subseteq X$ is $(t, |Y|^{-\beta})$-decomposable. Then for every $\epsilon \in (0, 1/8)$, there is a $t \cdot (2 + \epsilon)$-spanner for X with lightness $O_\epsilon\left(\frac{t}{\beta} \cdot n^\beta \cdot \log n\right)$ and sparsity $O_\epsilon\left(n^{1+\beta} \cdot \log n \cdot \log t\right)$.*

Proof. Using the same Algorithm 1, the analysis of the stretch and sparsity from Theorem 4 is still valid, since the number partitions taken in each scale is smaller than in Theorem 4. Recall that in scale i we set $\Delta_i = (1+\epsilon)^i$, and the size of the $\epsilon \cdot \Delta_i$-net N_i is $n_i \leq \max\{\frac{2L}{\epsilon \Delta_i}, n\}$. The difference from the previous proof is that N_i is $(t, n_i^{-\beta})$-decomposable, so the number of partitions taken is $\varphi_i = O(n_i^\beta \log n_i)$. In each partition we might add at most one edge per net point, and the weight of this edge is $O(t \cdot \Delta_i)$. We divide the edges of H to H_1 and H_2, and bound the weight of H_2 as above (using that $n_i \leq n$). For H_1 we get,

$$
\begin{aligned}
w(H_1) &\leq \sum_{i \in \left\{\log_{1+\epsilon} \frac{L}{n}, \ldots, \log_{1+\epsilon} L\right\}} O(t \cdot \Delta_i) \cdot n_i \cdot \varphi_i \\
&= O\left(t \cdot \sum_{i \in \left\{\log_{1+\epsilon} \frac{L}{n}, \ldots, \log_{1+\epsilon} L\right\}} \Delta_i \cdot \frac{L}{\epsilon \cdot \Delta_i} \cdot \left(\frac{L}{\epsilon \cdot \Delta_i}\right)^\beta \log \frac{L}{\epsilon \cdot \Delta_i}\right) \\
&= O_\epsilon\left(t \cdot \sum_{i \in \left\{\log_{1+\epsilon} \frac{L}{n}, \ldots, \log_{1+\epsilon} L\right\}} \left(\frac{L}{\Delta_i}\right)^\beta \cdot \log \frac{L}{\Delta_i}\right) \cdot L \\
&= O_\epsilon\left(t \cdot \sum_{i \in \left\{0, \ldots, \log_{1+\epsilon} n\right\}} (i+1) \cdot \left((1+\epsilon)^\beta\right)^i\right) \cdot L \,.
\end{aligned}
$$

Set the function $f(x) = \sum_{i=0}^{k} (i+1) \cdot x^i$, on the domain $(1, \infty)$, with parameter $k = \log_{1+\epsilon} n$. Then,

$$
\begin{aligned}
f(x) = \left(\int f dx\right)' = \left(\sum_{i=0}^{k} x^{i+1}\right)' = \left(\frac{x^{k+2} - x}{x - 1}\right)' \\
= \frac{\left((k+2)\, x^{k+1} - 1\right)(x-1) - \left(x^{k+2} - x\right)}{(x-1)^2} \leq \frac{(k+2)\, x^{k+1}}{x - 1} \,.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
w(H_1) &= O_\epsilon\left(t \cdot f\left((1+\epsilon)^\beta\right)\right) \cdot L \\
&= O_\epsilon\left(t \cdot \frac{\log_{1+\epsilon} n \cdot \left((1+\epsilon)^\beta\right)^{\log_{1+\epsilon} n}}{(1+\epsilon)^\beta - 1}\right) \cdot L = O_\epsilon\left(\frac{t}{\beta} \cdot n^\beta \cdot \log n\right) \cdot L \,.
\end{aligned}
$$

We conclude that the lightness of H is bounded by $O_\epsilon\left(\frac{t}{\beta} \cdot n^\beta \cdot \log n\right)$. ◀

In Section 5 we will show that any n-point Euclidean metric is $(t, n^{-O(1/t^2)})$-decomposable, and that for fixed $p \in (1,2)$, any n-point subset of ℓ_p is $(t, n^{-O(\log^2 t/t^p)})$-decomposable. The following corollaries are implied by Theorem 5 (rescaling t by a constant factor allows us to remove the $O(\cdot)$ term in the exponent of n, while obtaining stretch $O(t)$).

▶ **Corollary 6.** *For a set X of n points in Euclidean space, $t > 1$, there is an $O(t)$-spanner with lightness $O\left(t^3 \cdot n^{1/t^2} \cdot \log n\right)$ and $O\left(n^{1+1/t^2} \cdot \log n \cdot \log t\right)$ edges.*

▶ **Corollary 7.** *For a constant $p \in (1,2)$ and a set X of n points in ℓ_p space, there is an $O(t)$-spanner with lightness $O\left(\frac{t^{1+p}}{\log^2 t} \cdot n^{\log^2 t/t^p} \cdot \log n\right)$ and $O\left(n^{1+\log^2 t/t^p} \cdot \log n \cdot \log t\right)$ edges.*

▶ Remark. Corollary 6 applies for a set of points $X \subseteq \mathbb{R}^d$, where the dimension d is arbitrarily large. If $d = o(\log n)$ we can obtain improved spanners. Specifically, n-point subsets of d-dimensional Euclidean space are $(O(t), 2^{-d/t^2})$-decomposable (see Section 6). Applying Theorem 4 we obtain an $O(t)$-spanner with lightness $O_\epsilon\left(t \cdot 2^{d/t^2} \cdot \log^2 n\right)$ and $O_\epsilon\left(n \cdot 2^{d/t^2} \cdot \log n \cdot \log t\right)$ edges.

4.2 Doubling Metrics

It was shown in [1] that metrics with doubling constant λ are $(t, \lambda^{-O(1/t)})$-decomposable (the case $t = \Theta(\log \lambda)$ was given by [33]). Therefore, Theorem 4 implies:

▶ **Corollary 8.** *For every metric space (X, d_X) with doubling constant λ, and $t \geq 1$, there exist an $O(t)$-spanner with lightness $O\left(t \cdot \log^2 n \cdot \lambda^{1/t}\right)$ and $O\left(n \cdot \lambda^{1/t} \cdot \log n \cdot \log t\right)$ edges.*

4.3 Graph Spanners

In the case where the input is a graph G, it is natural to require that the spanner will be a *graph-spanner*, i.e., a subgraph of G. Given a (metric) spanner H, one can define a graph-spanner H' by replacing every edge $\{x,y\} \in H$ with the shortest path from x to y in G. It is straightforward to verify that the stretch and lightness of H' are no larger than those of H (however, the number of edges may increase).

Consider a graph G with genus g. In [3] it was shown that (the shortest path metric of) G is $\left(t, g^{-O(1/t)}\right)$-decomposable. Furthermore, graphs with genus g have $O(n+g)$ edges [32], so any graph-spanner will have at most so many edges. By Theorem 4 we have:

▶ **Corollary 9.** *Let G be a weighted graph on n vertices with genus g. Given a parameter $t \geq 1$, there exist an $O(t)$-graph-spanner of G with lightness $O\left(t \cdot \log^2 n \cdot g^{1/t}\right)$ and $O(n+g)$ edges.*

For general graphs, the transformation to graph-spanners described above may arbitrarily increase the number of edges (in fact, it will be bounded by $O(\sqrt{|E_H|} \cdot n)$, [20]). Nevertheless, if we have a *strong-decomposition*, we can modify Algorithm 1 to produce a sparse spanner. In a graph $G = (X, E)$, the *strong-diameter* of a cluster $A \subseteq X$ is $\max_{v,u \in A} d_{G[A]}(v,u)$, where $G[A]$ is the induced graph by A (as opposed to weak diameter, which is computed w.r.t the original metric distances). A partition $\mathcal{P}$ of X is Δ-*strongly-bounded* if the strong diameter of every $P \in \mathcal{P}$ is at most Δ. A distribution $\mathcal{D}$ over partitions of X is (t, Δ, δ)-*strong-decomposition*, if it is (t, Δ, δ)-decomposition and in addition every partition $\mathcal{P} \in \text{supp}(\mathcal{D})$ is Δ-strongly-bounded. A graph G is (t, δ)-*strongly-decomposable*, if for every $\Delta > 0$, the graph admits a $(\Delta, t \cdot \Delta, \delta)$-strong-decomposition.

▶ **Theorem 10.** *Let $G = (V, E, w)$ be a (t, δ)-strongly-decomposable, n-vertex graph with aspect ratio $\Lambda = \frac{\max_{e\in E} w(e)}{\min_{e\in E} w(e)}$. Then for every $\epsilon \in (0, 1)$, there is a $t \cdot (2 + \epsilon)$-graph-spanner for G with lightness $O_\epsilon\left(\frac{t}{\delta} \cdot \log^2 n\right)$ and $O_\epsilon(\frac{n}{\delta} \cdot \log n \cdot \log \Lambda)$ edges.*

Proof. We will execute Algorithm 1 with several modifications:

1. The for loop (in Line 2) will go over scales $i \in \{0, \dots, \log_{1+\epsilon} \Lambda\}$ (instead $\{0, \dots, \log_{1+\epsilon} L\}$).
2. We will use strong-decompositions instead of regular (weak) decompositions.
3. The partitions created in Line 3 will be over the set of all vertices V, rather then only net points N_i (as otherwise it will be impossible to get strong diameter).
   However, the requirement from close pairs to be clustered together (at least once), is still applied to net points only. Similarly to Claim 3, $\varphi_i = (2 \ln n_i)/\delta$ repetitions will suffice.
4. In Line 6, we will no longer add edges from v_P to all the net points in $P \in \mathcal{P}_j$. Instead, for every net point $x \in P \cap N_i$, we will add a shortest path in $G[P]$ from v_P to x. Note that all the edges added in all the clusters constitute a forest. Thus we add at most n edges per partition.

We now prove the stretch, sparsity and lightness of the resulting spanner.

Stretch. By the triangle inequality, it is enough to show small stretch guarantee only for edges (that is, only for $x, y \in V$ s.t. $\{x, y\} \in E$.) As we assumed that the minimal distance is 1, all the weights are within $[1, \Lambda]$. In particular, every edge $\{x, y\} \in E$ has weight $(1 + \epsilon)^{i-1} < w \le (1 + \epsilon)^i$ for $i \in \{0, \dots, \log_{1+\epsilon} \Lambda\}$. The rest of the analysis is similar to Theorem 4, with the only difference being that we use a path from v_P to $\tilde{x}$ rather than the edge $\{\tilde{x}, v_P\}$. This is fine since we only require that the length of this path is at most $(t \cdot (1 + 2\epsilon) \cdot \Delta)$, which is guaranteed by the strong diameter of clusters.

Sparsity. We have $O_\epsilon(\log \Lambda)$ scales. In each scale we had at most $\varphi_i \le \frac{2}{\delta} \log n$ partitions, where for each partition we added at most n edges. The bound on the sparsity follows.

Lightness. Consider scale i. We have n_i net points. For each net point we added at most one shortest path of weight at most $O(t \cdot \Delta_i)$ (as each cluster is $O(t \cdot \Delta_i)$-strongly bounded). As the number of partitions is φ_i, the total weight of all edges added at scale i is bounded by $O(t \cdot \Delta_i) \cdot n_i \cdot \varphi_i$. The rest of the analysis follows by similar lines to Theorem 4 (noting that $\Lambda < L$). ◀

5 LSH Induces Decompositions

In this section, we prove that LSH (locality sensitive hashing) induces decompositions. In particular, using the LSH schemes of [5, 46], we will get decompositions for ℓ_2 and ℓ_p spaces, $1 < p < 2$.

▶ **Definition 11.** (Locality-Sensitive-Hashing) Let H be a family of hash functions mapping a metric (X, d_X) to some universe U. We say that H is (r, cr, p_1, p_2)-sensitive if for every pair of points $x, y \in X$, the following properties are satisfied:

1. If $d_X(x, y) \le r$ then $\Pr_{h\in H}[h(x) = h(y)] \ge p_1$.
2. If $d_X(x, y) > cr$ then $\Pr_{h\in H}[h(x) = h(y)] \le p_2$.

Given an LSH, its parameter is $\gamma = \frac{\log 1/p_1}{\log 1/p_2}$. We will implicitly always assume that $p_1 \ge n^{-\gamma}$ ($n = |X|$), as indeed will occur in all the discussed settings. Andoni and Indyk [5] showed that for Euclidean space (ℓ_2), and large enough $t > 1$, there is an LSH with parameter

$\gamma = O\left(\frac{1}{t^2}\right)$. Nguyen [46], showed that for constant $p \in (1,2)$, and large enough $t > 1$, there is an LSH for ℓ_p, with parameter $\gamma = O\left(\frac{\log^2 t}{t^p}\right)$. We start with the following claim.

▶ **Claim 12.** Let (X, d_X) be a metric space, such that for every $r > 0$, there is an $(r, t \cdot r, p_1, p_2)$-sensitive LSH family with parameter γ. Then there is an $\left(r, t \cdot r, n^{-O(\gamma)}, n^{-2}\right)$-sensitive LSH family for X.

Proof. Set $k = \left\lceil \log_{\frac{1}{p_2}} n^2 \right\rceil \leq \frac{O(\log n)}{\log \frac{1}{p_2}}$, and let H be the promised $(r, t \cdot r, p_1, p_2)$-sensitive LSH family. We define an LSH family H' as follows. In order to sample $h \in H'$, pick $h_1, \dots, h_k$ uniformly and independently at random from H. The hash function h is defined as the concatenation of $h_1, \dots, h_k$. That is, $h(x) = (h_1(x), \dots, h_k(x))$.
For $x, y \in X$ such that $d_X(x, y) \geq t \cdot r$ it holds that

$$\Pr\left[h(x) = h(y)\right] = \Pi_i \Pr\left[h_i(x) = h_i(y)\right] \leq p_2^k \leq n^{-2} \ .$$

On the other hand, for $x, y \in X$ such that $d_X(x, y) \leq r$, it holds that

$$\Pr\left[h(x) = h(y)\right] = \Pi_i \Pr\left[h_i(x) = h_i(y)\right] \geq p_1^k = 2^{-\log \frac{1}{p_1} \cdot \frac{O(\log n)}{\log \frac{1}{p_2}}} = n^{-O(\gamma)} \ .$$

◀

▶ **Lemma 13.** *Let (X, d_X) be a metric space, such that for every $r > 0$, there is a $(r, t \cdot r, p_1, p_2)$-sensitive LSH family with parameter γ. Then (X, d_X) is $(t, n^{-O(\gamma)})$-decomposable.*

Proof. Let H' be an $\left(r, tr, n^{-O(\gamma)}, n^{-2}\right)$-sensitive LSH family, given by Claim 12. We will use H' in order to construct a decomposition for X. Each hash function $h \in H'$ induces a partition $\mathcal{P}_h$, by clustering all points with the same hash value, i.e. $\mathcal{P}_h(x) = \mathcal{P}_h(y) \iff h(x) = h(y)$. However, in order to ensure that our partition will be $t \cdot r$-bounded, we modify it slightly. For $x \in X$, if there is a $y \in \mathcal{P}_h(x)$ with $d_X(x, y) > t \cdot r$, remove x from $\mathcal{P}_h(x)$, and create a new cluster $\{x\}$. Denote by $\mathcal{P}'_h$ the resulting partition. $\mathcal{P}'_h$ is clearly $t \cdot r$-bounded, and we argue that every pair x, y at distance at most r is clustered together with probability at least $n^{-O(\gamma)}$. Denote by χ_x (resp., χ_y) the probability that x (resp., y) was removed from $\mathcal{P}_h(x)$ (resp., $\mathcal{P}_h(y)$). By the union bound on the at most n points in $\mathcal{P}_h(x)$, we have that both $\chi_x, \chi_y \leq 1/n$. We conclude

$$\Pr_{\mathcal{P}'_h}\left[\mathcal{P}'_h(x) = \mathcal{P}'_h(y)\right] \geq \Pr_{h \sim H}\left[h(x) = h(y)\right] - \Pr_h\left[\chi_x \vee \chi_y\right] \geq n^{-O(\gamma)} - \frac{2}{n} = n^{-O(\gamma)} \ .$$

◀

Using [5], Lemma 13 implies that ℓ_2 is $(t, n^{-O(1/t^2)})$-decomposable. Moreover, using [46] for constant $p \in (1,2)$, Lemma 13 implies that ℓ_p is $(t, n^{-O(\log^2 t / t^p)})$-decomposable.

6 Decomposition for d-Dimensional Euclidean Space

In Section 5, using a reduction from LSH, we showed that ℓ_2 is $(t, n^{-O(1/t^2)})$-decomposable. Here, we will show that for dimension $d = o(\log n)$, using a direct approach, better decomposition could be constructed.

Denote by $B_d(x, r)$ the d dimensional ball of radius r around x (w.r.t ℓ_2 norm). $V_d(r)$ denotes the volume of $B_d(x, r)$ (note that the center here is irrelevant). Denote by $C_d(u, r)$ the volume of the intersection of two balls of radius r, the centers of which are at distance u (i.e. for $\|x - y\|_2 = u$, $C_d(u, r)$ denotes the volume of $B_d(x, r) \cap B_d(y, r)$). We will use the following lemma which was proved in [5] (based on a lemma from [27]).

▶ **Lemma 14.** *([5]) For any $d \geq 2$ and $0 \leq u \leq r$*

$$\Omega\left(\frac{1}{\sqrt{d}}\right)\cdot\left(1-\left(\frac{u}{r}\right)^2\right)^{\frac{d}{2}} \leq \frac{C_d(u,r)}{V_d(r)} \leq \left(1-\left(\frac{u}{r}\right)^2\right)^{\frac{d}{2}} .$$

Using Lemma 14, we can construct better decompositions:

▶ **Lemma 15.** *For every $d \geq 2$ and $2 \leq t \leq \sqrt{2d/\ln d}$, ℓ_2^d is $O(t, 2^{-O(\frac{d}{t^2})})$-decomposable.*

Proof. Consider a set X of n points in ℓ_2^d, and fix $r > 0$. Let $\mathcal{B}$ be some box which includes all of X and such that each $x \in X$ is at distance at least $t \cdot r$ from the boundary of B. We sample points $s_1, s_2 \dots$ uniformly at random from $\mathcal{B}$. Set $P_i = B_X(s_i, \frac{t\cdot r}{2}) \setminus \bigcup_{j=1}^{i-1} B_X\left(s_j, \frac{t\cdot r}{2}\right)$. We sample points until $X = \bigcup_{i\geq 1} P_i$. Then, the partition will be $\mathcal{P} = \{P_1, P_2, \dots .\}$ (dropping empty clusters).

It is straightforward that $\mathcal{P}$ is $t\cdot r$-bounded. Thus it will be enough to prove that every pair x, y at distance at most r, has high enough probability to be clustered together. Let s_i be the first point sampled in $B_d\left(x, \frac{t\cdot r}{2}\right) \cup B_d\left(y, \frac{t\cdot r}{2}\right)$. By the minimality of i, $x, y \notin \bigcup_{j=1}^{i-1} B_d\left(s_j, \frac{t\cdot r}{2}\right)$ and thus both are yet un-clustered. If $s_i \in B_d\left(x, \frac{t\cdot r}{2}\right) \cap B_d\left(y, \frac{t\cdot r}{2}\right)$ then both x, y join P_i and thus clustered together. Using Lemma 14 we conclude,

$$\begin{aligned}
\Pr_{\mathcal{P}}[\mathcal{P}(x) = \mathcal{P}(y)] &= \Pr\Bigg[s_i \in B_d\left(x, \frac{t\cdot r}{2}\right) \cap B_d\left(y, \frac{t\cdot r}{2}\right) \\
&\qquad\qquad \Bigg| s_i \text{ is first in } B_d\left(x, \frac{t\cdot r}{2}\right) \cup B_d\left(y, \frac{t\cdot r}{2}\right)\Bigg] \\
&\geq \frac{C_d(\|x-y\|_2, \frac{t\cdot r}{2})}{2\cdot V_d(\frac{t\cdot r}{2})} \\
&= \Omega\left(\frac{1}{\sqrt{d}}\right)\left(1-\left(\frac{\|x-y\|_2}{\frac{t\cdot r}{2}}\right)^2\right)^{\frac{d}{2}} \\
&= \Omega\left(\frac{1}{\sqrt{d}}\right)\left(1-\frac{4}{t^2}\right)^{\frac{d}{2}} \\
&= \Omega\left(e^{-\frac{2d}{t^2}-\frac{1}{2}\ln d}\right) = 2^{-O(d/t^2)} .
\end{aligned}$$

◀

7 Acknowledgments

We would like to thank an anonymous reviewer for useful comments.

References

1 Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. *Advances in Mathematics*, 228(6):3026 – 3126, 2011. URL: `http://www.sciencedirect.com/science/article/pii/S000187081100288X`, `doi:10.1016/j.aim.2011.08.003`.

2 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, New Orleans, Louisiana, USA, January 7-10*, 2018.

3 Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: padded decomposition for minor-free graphs. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 79–88, 2014. URL: http://doi.acm.org/10.1145/2591796.2591849, doi:10.1145/2591796.2591849.

4 Ingo Althöfer, Gautam Das, David P. Dobkin, Deborah Joseph, and José Soares. On sparse spanners of weighted graphs. *Discrete & Computational Geometry*, 9:81–100, 1993. URL: https://doi.org/10.1007/BF02189308, doi:10.1007/BF02189308.

5 Alexandr Andoni and Piotr Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. In *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 21-24 October 2006, Berkeley, California, USA, Proceedings*, pages 459–468, 2006. URL: https://doi.org/10.1109/FOCS.2006.49, doi:10.1109/FOCS.2006.49.

6 Baruch Awerbuch. Communication-time trade-offs in network synchronization. In *Proc. of 4th PODC*, pages 272–276, 1985.

7 Baruch Awerbuch. Complexity of network synchronization. *J. ACM*, 32(4):804–823, October 1985. URL: http://doi.acm.org/10.1145/4221.4227, doi:10.1145/4221.4227.

8 Baruch Awerbuch, Alan E. Baratz, and David Peleg. Cost-sensitive analysis of communication protocols. In *Proceedings of the Ninth Annual ACM Symposium on Principles of Distributed Computing, Quebec City, Quebec, Canada, August 22-24, 1990*, pages 177–187, 1990. URL: http://doi.acm.org/10.1145/93385.93417, doi:10.1145/93385.93417.

9 Baruch Awerbuch, Alan E. Baratz, and David Peleg. Efficient broadcast and light-weight spanners. *Manuscript*, 1991.

10 Y. Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *Proc. of 37th FOCS*, pages 184–193, 1996.

11 Punyashloka Biswal, James R. Lee, and Satish Rao. Eigenvalue bounds, spectral partitioning, and metrical deformations via flows. *J. ACM*, 57(3), 2010.

12 Glencora Borradaile, Hung Le, and Christian Wulff-Nilsen. Greedy spanners are optimal in doubling metrics. *CoRR*, abs/1712.05007, 2017. URL: http://arxiv.org/abs/1712.05007, arXiv:1712.05007.

13 Glencora Borradaile, Hung Le, and Christian Wulff-Nilsen. Minor-free graphs have light spanners. In *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*, pages 767–778, 2017. URL: https://doi.org/10.1109/FOCS.2017.76, doi:10.1109/FOCS.2017.76.

14 R. Braynard, D. Kostic, A. Rodriguez, J. Chase, and A. Vahdat. Opus: an overlay peer utility service. In *Prof. of 5th OPENARCH*, 2002.

15 Gruia Calinescu, Howard Karloff, and Yuval Rabani. Approximation algorithms for the 0-extension problem. *SIAM J. Comput.*, 34(2):358–372, February 2005. URL: http://dx.doi.org/10.1137/S0097539701395978, doi:10.1137/S0097539701395978.

16 P. B. Callahan and S. R. Kosaraju. A decomposition of multi-dimensional point-sets with applications to k-nearest-neighbors and n-body potential fields. In *Proc. of 24th STOC*, pages 546–556, 1992.

17 B. Chandra, G. Das, G. Narasimhan, and J. Soares. New sparseness results on graph spanners. *Int. J. Comput. Geometry Appl.*, 5:125–144, 1995.

18 Shiri Chechik and Christian Wulff-Nilsen. Near-optimal light spanners. In *Proc. of 27th SODA*, pages 883–892, 2016.

19 Edith Cohen. Fast algorithms for constructing t-spanners and paths with stretch t. *SIAM J. Comput.*, 28(1):210–236, 1998. URL: https://doi.org/10.1137/S0097539794261295, doi:10.1137/S0097539794261295.

20 Don Coppersmith and Michael Elkin. Sparse sourcewise and pairwise distance preservers. *SIAM J. Discrete Math.*, 20(2):463–501, 2006. URL: `https://doi.org/10.1137/050630696`, `doi:10.1137/050630696`.

21 Gautam Das, Paul J. Heffernan, and Giri Narasimhan. Optimally sparse spanners in 3-dimensional euclidean space. In *Proceedings of the Ninth Annual Symposium on Computational GeometrySan Diego, CA, USA, May 19-21, 1993*, pages 53–62, 1993. URL: `http://doi.acm.org/10.1145/160985.160998`, `doi:10.1145/160985.160998`.

22 Amin Vahdat Dejan Kostic. Latency versus cost optimizations in hierarchical overlay networks. *Technical report, Duke University*, (CS-2001-04), 2002.

23 Michael Elkin. Computing almost shortest paths. *ACM Trans. Algorithms*, 1(2):283–323, 2005. URL: `http://doi.acm.org/10.1145/1103963.1103968`, `doi:10.1145/1103963.1103968`.

24 Michael Elkin, Ofer Neiman, and Shay Solomon. Light spanners. In *Proc. of 41th ICALP*, pages 442–452, 2014.

25 Michael Elkin and Jian Zhang. Efficient algorithms for constructing (1+epsilon, beta)-spanners in the distributed and streaming models. *Distributed Computing*, 18(5):375–385, 2006. URL: `https://doi.org/10.1007/s00446-005-0147-2`, `doi:10.1007/s00446-005-0147-2`.

26 Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, STOC '03, pages 448–455, New York, NY, USA, 2003. ACM. URL: `http://doi.acm.org/10.1145/780542.780608`, `doi:http://doi.acm.org/10.1145/780542.780608`.

27 Uriel Feige and Gideon Schechtman. On the optimality of the random hyperplane rounding technique for max cut. *Random Struct. Algorithms*, 20(3):403–440, May 2002. URL: `http://dx.doi.org/10.1002/rsa.10036`, `doi:10.1002/rsa.10036`.

28 Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. Graph distances in the streaming model: the value of space. In *Proc. of 16th SODA*, pages 745–754, 2005.

29 Arnold Filtser and Shay Solomon. The greedy spanner is existentially optimal. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing, PODC 2016, Chicago, IL, USA, July 25-28, 2016*, pages 9–17, 2016. URL: `http://doi.acm.org/10.1145/2933057.2933114`, `doi:10.1145/2933057.2933114`.

30 Lee-Ad Gottlieb. A light metric spanner. In *Proc. of 56th FOCS*, pages 759–772, 2015.

31 Michelangelo Grigni. Approximate TSP in graphs with forbidden minors. In *Proc. of 27th ICALP*, pages 869–877, 2000.

32 Jonathan L. Gross and Thomas W. Tucker. *Topological Graph Theory*. Wiley-Interscience, New York, NY, USA, 1987.

33 Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proc. of 44th FOCS*, pages 534–543, 2003.

34 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, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 804–809, 2013. URL: `https://doi.org/10.1137/1.9781611973105.57`, `doi:10.1137/1.9781611973105.57`.

35 Sariel Har-Peled and Manor Mendel. Fast construction of nets in low-dimensional metrics and their applications. *SIAM J. Comput.*, 35(5):1148–1184, 2006. URL: `https://doi.org/10.1137/S0097539704446281`, `doi:10.1137/S0097539704446281`.

36 Jonathan A. Kelner, James R. Lee, Gregory N. Price, and Shang-Hua Teng. Higher eigenvalues of graphs. In *FOCS*, pages 735–744, 2009.

37 Philip N. Klein, Serge A. Plotkin, and Satish Rao. Excluded minors, network decomposition, and multicommodity flow. In *STOC*, pages 682–690, 1993.

38 Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: A new embedding method for finite metrics. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 434–443, Washington, DC, USA, 2004. IEEE Computer Society. URL: `http://portal.acm.org/citation.cfm?id=1032645.1033199`, `doi:10.1109/FOCS.2004.41`.

39 J. R. Lee and A. Naor. Extending lipschitz functions via random metric partitions. *Inventiones Mathematicae*, 160(1):59–95, 2005.

40 James R. Lee and Anastasios Sidiropoulos. Genus and the geometry of the cut graph. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2010, Austin, Texas, USA, January 17-19, 2010*, pages 193–201, 2010. URL: `https://doi.org/10.1137/1.9781611973075.18`, `doi:10.1137/1.9781611973075.18`.

41 Tom Leighton and Satish Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM*, 46:787–832, November 1999. URL: `http://doi.acm.org/10.1145/331524.331526`, `doi:http://doi.acm.org/10.1145/331524.331526`.

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

43 Nathan Linial and Michael Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993. (Preliminary version in *2nd SODA*, 1991).

44 Manor Mendel and Assaf Naor. Ramsey partitions and proximity data structures. *Journal of the European Mathematical Society*, 9(2):253–275, 2007.

45 Giri Narasimhan and Michiel H. M. Smid. *Geometric spanner networks*. Cambridge University Press, 2007.

46 Huy L. Nguyen. Approximate nearest neighbor search in ℓ_p. *CoRR*, abs/1306.3601, 2013. URL: `http://arxiv.org/abs/1306.3601`, `arXiv:1306.3601`.

47 David Peleg. Proximity-preserving labeling schemes and their applications. In *Graph-Theoretic Concepts in Computer Science, 25th International Workshop, WG '99, Ascona, Switzerland, June 17-19, 1999, Proceedings*, pages 30–41, 1999. URL: `https://doi.org/10.1007/3-540-46784-X_5`, `doi:10.1007/3-540-46784-X_5`.

48 David Peleg. *Distributed Computing: A Locality-Sensitive Approach*. SIAM, Philadelphia, PA, 2000.

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

50 David Peleg and Eli Upfal. A trade-off between space and efficiency for routing tables. *J. ACM*, 36(3):510–530, 1989.

51 Satish B. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *SOCG*, pages 300–306, 1999.

52 L. Roditty and U. Zwick. On dynamic shortest paths problems. In *Proc. of 32nd ESA*, pages 580–591, 2004.

53 Liam Roditty, Mikkel Thorup, and Uri Zwick. Deterministic constructions of approximate distance oracles and spanners. In *Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings*, pages 261–272, 2005. URL: `https://doi.org/10.1007/11523468_22`, `doi:10.1007/11523468_22`.

54 J. S. Salowe. Construction of multidimensional spanner graphs, with applications to minimum spanning trees. In *Proc. of 7th SoCG*, pages 256–261, 1991.

55 Michiel H. M. Smid. The weak gap property in metric spaces of bounded doubling dimension. In *Efficient Algorithms, Essays Dedicated to Kurt Mehlhorn on the Occasion of His 60th Birthday*, pages 275–289, 2009. URL: `https://doi.org/10.1007/978-3-642-03456-5_19`, `doi:10.1007/978-3-642-03456-5_19`.

56 Kunal Talwar. Bypassing the embedding: algorithms for low dimensional metrics. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 281–290, 2004. URL: `http://doi.acm.org/10.1145/1007352.1007399`, `doi:10.1145/1007352.1007399`.

57 Mikkel Thorup and Uri Zwick. Compact routing schemes. In *Proc. of 13th SPAA*, pages 1–10, 2001.

58 Mikkel Thorup and Uri Zwick. Approximate distance oracles. *J. ACM*, 52(1):1–24, 2005. URL: `http://doi.acm.org/10.1145/1044731.1044732`, `doi:10.1145/1044731.1044732`.

59 P. M. Vaidya. A sparse graph almost as good as the complete graph on points in k dimensions. *Discrete & Computational Geometry*, 6:369–381, 1991.

60 Jürgen Vogel, Jörg Widmer, Dirk Farin, Martin Mauve, and Wolfgang Effelsberg. Priority-based distribution trees for application-level multicast. In *Proceedings of the 2nd Workshop on Network and System Support for Games, NETGAMES 2003, Redwood City, California, USA, May 22-23, 2003*, pages 148–157, 2003. URL: `http://doi.acm.org/10.1145/963900.963914`, `doi:10.1145/963900.963914`.

61 Bang Ye Wu, Kun-Mao Chao, and Chuan Yi Tang. Light graphs with small routing cost. *Networks*, 39(3):130–138, 2002. URL: `http://dx.doi.org/10.1002/net.10019`, `doi:10.1002/net.10019`.