Sparse Euclidean Spanners with Tiny Diameter: A Tight Lower Bound

Hung Le ✉
University of Massachusetts, Amherst, MA, USA

Lazar Milenković ✉
Tel Aviv University, Israel

Shay Solomon ✉
Tel Aviv University, Israel

Abstract

In STOC'95 [ADMSS95] Arya et al. showed that any set of n points in $\mathbb{R}$ admits a $(1+\epsilon)$-spanner with hop-diameter at most 2 (respectively, 3) and $O(n \log n)$ edges (resp., $O(n \log\log n)$ edges). They also gave a general upper bound tradeoff of hop-diameter at most k and $O(n\alpha_k(n))$ edges, for any $k \geq 2$. The function α_k is the inverse of a certain Ackermann-style function at the $\lfloor k/2 \rfloor$th level of the primitive recursive hierarchy, where $\alpha_0(n) = \lceil n/2 \rceil$, $\alpha_1(n) = \lceil \sqrt{n} \rceil$, $\alpha_2(n) = \lceil \log n \rceil$, $\alpha_3(n) = \lceil \log\log n \rceil$, $\alpha_4(n) = \log^* n$, $\alpha_5(n) = \lfloor \frac{1}{2} \log^* n \rfloor$, …. Roughly speaking, for $k \geq 2$ the function α_k is close to $\lfloor \frac{k-2}{2} \rfloor$-iterated log-star function, i.e., log with $\lfloor \frac{k-2}{2} \rfloor$ stars. Also, $\alpha_{2\alpha(n)+4}(n) \leq 4$, where $\alpha(n)$ is the one-parameter inverse Ackermann function, which is an extremely slowly growing function.

Whether or not this tradeoff is tight has remained open, even for the cases $k = 2$ and $k = 3$. Two lower bounds are known: The first applies only to spanners with stretch 1 and the second is sub-optimal and applies only to sufficiently large (constant) values of k. In this paper we prove a tight lower bound for any constant k: For any fixed $\epsilon > 0$, any $(1+\epsilon)$-spanner for the uniform line metric with hop-diameter at most k must have at least $\Omega(n\alpha_k(n))$ edges.

2012 ACM Subject Classification Theory of computation → Computational geometry

Keywords and phrases Euclidean spanners, hop-diameter, inverse-Ackermann, lower bounds, sparse spanners

Digital Object Identifier 10.4230/LIPIcs.SoCG.2022.54

Related Version *Full Version*: https://arxiv.org/abs/2112.09124

Funding *Hung Le*: Supported by National Science Foundation under Grant No. CCF-2121952.
Lazar Milenković: Partially supported by the Israel Science Foundation (ISF) grant No.1991/1, and by a grant from the United States-Israel Binational Science Foundation (BSF), Israel, and the United States National Science Foundation (NSF).
Shay Solomon: Supported by the Israel Science Foundation (ISF) grant No.1991/1, and by a grant from the United States-Israel Binational Science Foundation (BSF), Israel, and the United States National Science Foundation (NSF).

1 Introduction

Consider a set S of n points in $\mathbb{R}^d$ and a real number $t \geq 1$. A weighted graph $G = (S, E, w)$ in which the weight function is given by the Euclidean distance, i.e., $w(x, y) = \|x - y\|$ for each $e = (x, y) \in E$, is called a *geometric graph*. We say that a geometric graph G is a *t-spanner* for S if for every pair $p, q \in S$ of distinct points, there is a path in G between p and q whose *weight* (i.e., the sum of all edge weights in it) is at most t times the Euclidean distance $\|p - q\|$ between p and q. Such a path is called a *t-spanner path*. The problem of constructing Euclidean spanners has been studied intensively over the years

[15, 25, 4, 10, 16, 5, 17, 32, 2, 11, 18, 35, 34, 19, 27]. Euclidean spanners are of importance both in theory and in practice, as they enable approximation of the complete Euclidean graph in a more succinct form; in particular, they find a plethora of applications, e.g., in geometric approximation algorithms, network topology design, geometric distance oracles, distributed systems, design of parallel machines, and other areas [16, 28, 32, 20, 22, 21, 23, 29]. We refer the reader to the book by Narasimhan and Smid [30], which provides a thorough account on Euclidean spanners and their applications.

In terms of applications, the most basic requirement from a spanner (besides achieving a small stretch) is to be *sparse*, i.e., to have only a small number of edges. However, for many applications, the spanner is required to preserve some additional properties of the underlying complete graph. One such property, which plays a key role in various applications (such as to routing protocols) [6, 1, 2, 11, 18, 24], is the *hop-diameter*: a t-spanner for S is said to have an hop-diameter of k if, for any $p, q \in S$, there is a t-spanner path between p and q with at most k edges (or hops).

1.1 Known upper bounds

1-spanners for tree metrics. We denote the tree metric induced by an n-vertex (possibly weighted) rooted tree (T, rt) by M_T. A spanning subgraph G of M_T is said to be a *1-spanner* for T, if for every pair of vertices, their distance in G is equal to their distance in T. The problem of constructing 1-spanners for tree metrics is a fundamental one, and has been studied quite extensively over the years, also in more general settings, such as planar metrics [38], general metrics [37] and general graphs [8]. This problem is also intimately related to the extremely well-studied problems of computing partial-sums and online product queries in semigroup and their variants (see [36, 39, 3, 13, 31, 2], and the references therein).

Alon and Schieber [3] and Bodlaender et al. [9] showed that for any n-point tree metric, a 1-spanner with diameter 2 (respectively, 3) and $O(n \log n)$ edges (resp., $O(n \log \log n)$ edges) can be built within time linear in its size. For $k \geq 4$, Alon and Schieber [3] showed that 1-spanners with diameter at most $2k$ and $O(n\alpha_k(n))$ edges can be built in $O(n\alpha_k(n))$ time. The function α_k is the inverse of a certain Ackermann-style function at the $\lfloor k/2 \rfloor$th level of the primitive recursive hierarchy, where $\alpha_0(n) = \lceil n/2 \rceil$, $\alpha_1(n) = \lceil \sqrt{n} \rceil$, $\alpha_2(n) = \lceil \log n \rceil$, $\alpha_3(n) = \lceil \log \log n \rceil$, $\alpha_4(n) = \log^* n$, $\alpha_5(n) = \left\lfloor \frac{1}{2} \log^* n \right\rfloor$, etc. Roughly speaking, for $k \geq 2$ the function α_k is close to $\lfloor \frac{k-2}{2} \rfloor$-iterated log-star function, i.e., log with $\lfloor \frac{k-2}{2} \rfloor$ stars. Also, $\alpha_{2\alpha(n)+2}(n) \leq 4$, where $\alpha(n)$ is the one-parameter inverse Ackermann function, which is an extremely slowly growing function. (The functions $\alpha_k(n)$ and $\alpha(n)$ are formally defined in [3, 33]; see also Section 2 of the full version [26].) Bodlaender et al. [9] constructed 1-spanners with diameter at most k and $O(n\alpha_k(n))$ edges, with a high running time. Solomon [33] gave a construction that achieved the best of both worlds: a tradeoff of k versus $O(n\alpha_k(n))$ between the hop-diameter and the number of edges in linear time of $O(n\alpha_k(n))$.

Alternative constructions, given by Yao [39] for line metrics and later extended by Chazelle [12] to general tree metrics, achieve a tradeoff of m edges versus $\Theta(\alpha(m, n))$ hop-diameter, where $\alpha(m, n)$ is the standard two-parameter inverse Ackermann function [36]; see also Section 2 of the full version [26]. However, these constructions provide 1-spanners with diameter $\Gamma' \cdot k$ rather than $2k$ or k, for some constant $\Gamma' > 30$.

$(1+\epsilon)$-spanners. In STOC'95 Arya et al. [5] proved the so-called "Dumbbell Theorem", which states that, for any d-dimensional Euclidean space, a $(1+\epsilon, O(\frac{\log(1/\epsilon)}{\epsilon^d}))$-tree cover can be constructed in $O(\frac{\log(1/\epsilon)}{\epsilon^d} \cdot n \log n + \frac{1}{\epsilon^{2d}} \cdot n)$ time; see Section 2 for the definition of tree cover. The Dummbell Theorem implies that any construction of 1-spanners for tree metrics can be translated into a construction of Euclidean $(1+\epsilon)$-spanners. Applying the construction of 1-spanners for tree metrics from [33], this gives rise to an optimal $O(n \log n)$-time construction (in the algebraic computation tree (ACT) model[1]) of Euclidean $(1+\epsilon)$-spanners. This result can be generalized (albeit not in the ACT model) for the wider family of *doubling metrics*, by using the tree cover theorem of Bartal et al. [7], which generalizes the Dumbbell Theorem of [5] for arbitrary doubling metrics.

1.2 Known lower bounds

The first lower bound on 1-spanners for tree metrics was given by Yao [39] and it establishes a tradeoff of m edges versus hop-diameter of $\Omega(\alpha(m,n))$ for the uniform line metric. Alon and Schieber [3] gave a stronger lower bound on 1-spanners for the uniform line metric: hop-diameter k versus $\Omega(n\alpha_k(n))$ edges, for any k; it is easily shown that this lower bound implies that of [39] (see Appendix A of the full version [26]), but the converse is not true.

The above lower bounds apply to 1-spanners. There is also a lower bound on $(1+\epsilon)$-spanners that applies to line metrics, by Chan and Gupta [11], which extends that of [39]: m edges versus hop-diameter of $\Omega(\alpha(m,n))$. As mentioned already concerning this tradeoff, it only provides a meaningful bound for sufficiently large values of hop-diameter (above say 30), and it does not apply to hop-diameter values that approach 1, which is the focus of this work. More specifically, it can be used to show that any $(1+\epsilon)$-spanner for a certain line metric with hop-diameter at most k must have $\Omega(n\alpha_{2k+6}(n))$ edges. When $k=2$ (resp. $k=3$), this gives $\Omega(n \log^{****} n)$ (resp. $\Omega(n \log^{*****} n)$) edges, which is far from the upper bound of $O(n \log n)$ (resp., $O(n \log\log n)$). Furthermore, the line metric used in the proof of [11] is not as basic as the uniform line metric – it is derived from hierarchically well-separated trees (HSTs), and to achieve the result for line metrics, an embedding from HSTs to the line with an appropriate separation parameter is employed. The resulting line metric is very far from a uniform one and its aspect ratio[2] depends on the stretch – it will be super-polynomial whenever ϵ is sufficiently small or sufficiently large; of course, the aspect ratio of the uniform line metric (which is the metric used by [39, 3]) is linear in n. As point sets arising in real-life applications (e.g., for various random distributions) have polynomially bounded aspect ratio, it is natural to ask whether one can achieve a lower bound for a point set of polynomial aspect ratio.

1.3 Our contribution

We prove that any $(1+\epsilon)$-spanner for the uniform line metric with hop-diameter k must have at least $\Omega(n\alpha_k(n))$ edges, for any constant $k \geq 2$.

▶ **Theorem 1.** *For any positive integer n, any integer $k \geq 2$ and any $\epsilon \in [0, 1/2]$, any $(1+\epsilon)$-spanner with hop-diameter k for the uniform line metric with n points must contain at least $\Omega(\frac{n}{2^{6\lfloor k/2 \rfloor}}\alpha_k(n))$ edges.*

Interestingly, our lower bound applies also to any $\epsilon > 1/2$, where the bound on the number of edges reduces linearly with ϵ, i.e., it becomes $\Omega(n\alpha_k(n)/\epsilon)$. We stress that our lower bound instance, namely the uniform line metric, does not depend on ϵ, and the lower bound that it provides holds *simultaneously for all values of* ϵ.

[1] Refer to Chapter 3 in [30] for the definition of the ACT model. A matching lower bound of $\Omega(n \log n)$ on the time needed to construct Euclidean spanners is given in [14].

[2] The *aspect ratio* of a metric is the ratio of the maximum pairwise distance to the minimum one.

Although our lower bound on the number of edges coincides with $\Omega(n\alpha_k(n))$ only for constant k, we note that the values of k of interest range between 1 and $O(\alpha(n))$, where $\alpha(\cdot)$ is a very slowly growing function, e.g., $\alpha(n)$ is asymptotically much smaller than $\log^* n$. Indeed, as mentioned, for $k = 2\alpha(n) + 4$, we have $\alpha_{2\alpha(n)+4}(n) \leq 4$, and clearly any spanner must have $\Omega(n)$ edges. Thus the gap between our lower bound on the number of edges and $\Omega(n\alpha_k(n))$, namely, a multiplicative factor of $2^{6\lfloor k/2 \rfloor}$, which in particular is no greater than $2^{O(\alpha(n))}$, is very small.

For technical reasons we prove a more general lower bound, stated in Theorem 17. In particular, we need to consider a more general notion of Steiner spanners[3], and to prove the lower bound for a certain family of line metrics to which the uniform line metric belongs; Theorem 1 follows directly from Theorem 17. See Section 2 for the definitions.

For constant values of k, Theorem 1 strengthens the lower bound shown by [3], which applies only to stretch 1, whereas our tradeoff holds for arbitrary stretch. Whether or not the term $\frac{1}{2^{6\lfloor k/2 \rfloor}}$ in the bound on the number of edges in Theorem 1 can be removed is left open by our work. As mentioned before, we show in Appendix A of the full version [26] that this tradeoff implies the tradeoff by [39] (for stretch 1) and [11] (for larger stretch).

The proof overview appears in the full version [26].

2 Preliminaries

▶ **Definition 2** (Tree covers). *Let $M_X = (X, \delta_X)$ be an arbitrary metric space. We say that a weighted tree T is a* dominating *tree for M_X if $X \subseteq V(T)$ and it holds that $\delta_T(x, y) \geq \delta_X(x, y)$, for every $x, y \in X$. For $\gamma \geq 1$ and an integer $\zeta \geq 1$, a (γ, ζ)-*tree cover *of $M_X = (X, \delta_X)$ is a collection of ζ* dominating trees *for M_X, such that for every $x, y \in X$, there exists a tree T with $d_T(u, v) \leq \gamma \cdot \delta_X(u, v)$; we say that the* stretch *between x and y in T is at most γ, and the parameter γ is referred to as the* stretch *of the tree cover.*

▶ **Definition 3** (Uniform line metric). *A uniform line metric $U = (\mathbb{Z}, d)$ is a metric on a set of integer points such that the distance between two points $a, b \in \mathbb{Z}$, denoted by $d(a, b)$ is their Euclidean distance, which is $|a - b|$. For two integers $l, r \in \mathbb{Z}$, such that $l \leq r$, we define a uniform line metric on an interval $[l, r]$, denoted by $U(l, r)$, as a subspace of U consisting of all the integer points k, such that $l \leq k \leq r$. We use $U(n)$ to denote a uniform line metric on the interval $[1, n]$.*

Although we aim to prove the lower bound for uniform line metric, the inductive nature of our argument requires several generalizations of the considered metric space and spanner.

▶ **Definition 4** (t-sparse line metric). *Let l and r be two integers such that $l < r$. We call metric space $U((l, r), t)$ t-sparse if:*

- *It is a subspace of $U(l, r)$.*
- *Each of the consecutive intervals of $[l, r]$ of size t ($[l, l + t - 1], [l + t, l + 2t - 1], \ldots$) contains exactly one point. These intervals are called $((l, r), t)$-intervals and the point inside each such interval is called* representative *of the interval.*

▶ Remark 5. Throughout the paper, we will always consider Steiner spanners that can contain arbitrary points from the uniform line metric.

[3] A Steiner spanner for a point set S is a spanner that may contain additional Steiner points (which do not belong to S). Clearly, a lower bound for Steiner spanners also applies to ordinary spanners.

▶ **Definition 6** (Global hop-diameter). *For any two integers l, r such that $r = l + nt - 1$, let $U((l,r),t)$ be a t-sparse line metric with n points and let X be a subspace of $U((l,r),t)$. An edge that connects two points is* $((l,r),t)$-global *if it has endpoints in two different $((l,r),t)$-intervals of $U((l,r),t)$. A spanner on X with stretch $(1+\epsilon)$ has its* $((l,r),t)$-global hop-diameter *bounded by k if every pair of points in X has a path of stretch at most $(1+\epsilon)$ consisting of at most k $((l,r),t)$-global edges.*

For ease of presentation, we focus on $\epsilon \in [0, 1/2]$, as this is the basic regime. Our argument naturally extends to any $\epsilon > 1/2$, with the lower bound degrading by a factor of $1/\epsilon$.

▶ **Lemma 7** (Separation property). *Let $l, r, t \in \mathbb{N}$, $l \leq r$, $t \geq 1$ and let $i \coloneqq \lceil \frac{1+\epsilon/2}{1+\epsilon} l + \frac{\epsilon/2}{1+\epsilon} r \rceil$, and $j \coloneqq \lfloor \frac{\epsilon/2}{1+\epsilon} l + \frac{1+\epsilon/2}{1+\epsilon} r \rfloor$. Let a, b be two points in $U((l,r),t)$ such that $i \leq a < b \leq j$. Then, any $(1+\epsilon)$-spanner path between a and b contains points strictly inside $[l, r]$.*

▶ **Corollary 8.** *For every integer $N \geq 34$ and any t-sparse line metric $U((1,N),t)$, any spanner path with stretch at most $3/2$ between metric points a and b such that $\lfloor N/4 \rfloor \leq a \leq b \leq \lceil 3N/4 \rceil$ contains points strictly inside $[1, N]$.*

3 Warm-up: lower bounds for hop-diameters 2 and 3

In this section, we prove the lower bound for cases $k = 2$ (Lemma 10 in Section 3.1) and $k = 3$ (Lemma 13 in Section 3.2). In fact, we prove more general statements (Theorems 9 and 12), which apply not only to uniform line metric, but to subspaces of t-sparse line metrics, where a constant fraction of the points is missing. We use these general statements in Section 4, to prove the result for general k (cf. Theorem 17).

3.1 Hop diameter 2

▶ **Theorem 9.** *For any two positive integers $n \geq 1000$ and t, and any two integers l, r such that $r = l + nt - 1$, let $U((l,r),t)$ be a t-sparse line metric with n points and let X be a subspace of $U((l,r),t)$ which contains at least $\frac{31}{32}n$ points. Then, for any choice of $\epsilon \in [0, 1/2]$, any spanner on X with $((l,r),t)$-global hop-diameter 2 and stretch $1+\epsilon$ contains at least $T'_2(n) \geq \frac{n}{256} \cdot \alpha_2(n)$ $((l,r),t)$-global edges which have both endpoints inside $[l, r]$.*

The theorem is proved in three steps. First, we prove Lemma 10, which concerns uniform line metrics. Then, we prove Lemma 11 for a subspace that contains at least $31/32$ fraction of the points of the original metric. In the third step, we observe that the same argument applies for t-sparse line metrics.

▶ **Lemma 10.** *For any positive integer n, and any two integers l, r such that $r = l + n - 1$, let $U(l,r)$ be a uniform line metric with n points. Then, for any choice of $\epsilon \in [0, 1/2]$, any spanner on $U(l,r)$ with hop-diameter 2 and stretch $1+\epsilon$ contains at least $T_2(n) \geq \frac{1}{16} \cdot n \log n$ edges which have both endpoints inside $[l, r]$.*

Proof. Suppose without loss of generality that we are working on the uniform line metric $U(1,n)$. Let H be an arbitrary $(1+\epsilon)$-spanner for $U(1,n)$ with hop-diameter 2.

For the base case, we take $64 \leq n \leq 127$. In that case our lower bound is $\frac{n}{16} \cdot \log n < n - 1$, which is a trivial lower bound for the number of edges in H, since every two consecutive points must be connected via a direct edge.

For the proof of the inductive step, we can assume that $n \geq 128$. We would like to prove that the number of spanner edges in H is lower bounded by $T_2(n)$, which satisfies recurrence $T_2(n) = 2T_2(\lfloor n/2 \rfloor) + 11n/64$ with the base case $T_2(n) = (n/16)\log n$ when $n \leq 128$. Split the interval into two disjoint parts: the left part $[1, \lfloor n/2 \rfloor]$ and the right part $[\lfloor n/2 \rfloor + 1, n]$. From the induction hypothesis on the uniform line metric $U(1, \lfloor n/2 \rfloor)$ we know that any spanner with hop-diameter 2 and stretch $1 + \epsilon$ contains at least $T_2(\lfloor n/2 \rfloor)$ edges that have both endpoints inside $[1, \lfloor n/2 \rfloor]$. Similarly, any spanner for $U(\lfloor n/2 \rfloor + 1, n)$ contains at least $T_2(\lfloor n/2 \rfloor)$ edges that have both endpoints inside $[\lfloor n/2 \rfloor + 1, n]$. This means that the sets of edges considered on the left side and the right side are disjoint. We will show below that there are $\Omega(n)$ edges that have one point on the left and the other on the right.

Consider the set L, consisting of the points inside $[n/4, \lfloor n/2 \rfloor]$ and the set R, consisting of the points in $[\lfloor n/2 \rfloor + 1, 3n/4]$. From Corollary 8, since n is sufficiently large, we know that any $(1 + \epsilon)$-spanner path connecting point $a \in L$ and $b \in R$ has to have all its points inside $[1, n]$. We use term *cross edge* to denote any edge that has one endpoint in the left part and the other endpoint in the right part. We claim that any spanner with hop-diameter at most 2 and stretch $1 + \epsilon$ has to contain at least $\min(|L|, |R|)$ cross edges. For this particular choice of $|L|$ and $|R|$, we have that $\min(|L|, |R|) = |R|$. Suppose for contradiction that the spanner contains less than $|R|$ cross edges. This means that at least one point in $x \in R$ is not connected via a direct edge to any point on the left. Observe that, for every point $l \in L$, the 2-hop spanner path between x and l must be of the form (x, r_l, l) for some point r_l in the right set. It follows that every $l \in L$ induces a different cross edge (r_l, l). Thus, the number of cross edges, denoted by $|E_C|$, is $|R| \geq |L|$, which is a contradiction. From the definition of L and R, we know that $\min(|L|, |R|) \geq n/4 - 2$, implying that the number of cross edges is at least $n/4 - 2 \geq 11n/64$, for all $n \geq 26$. (See also Figure 1 for an illustration.) Thus, we have: $T_2(n) = 2T_2(\lfloor n/2 \rfloor) + \frac{11n}{64} \geq 2 \cdot \frac{\lfloor n/2 \rfloor}{16} \log \lfloor n/2 \rfloor + \frac{11n}{64} \geq \frac{n}{16} \cdot \log n$ as claimed. ◀

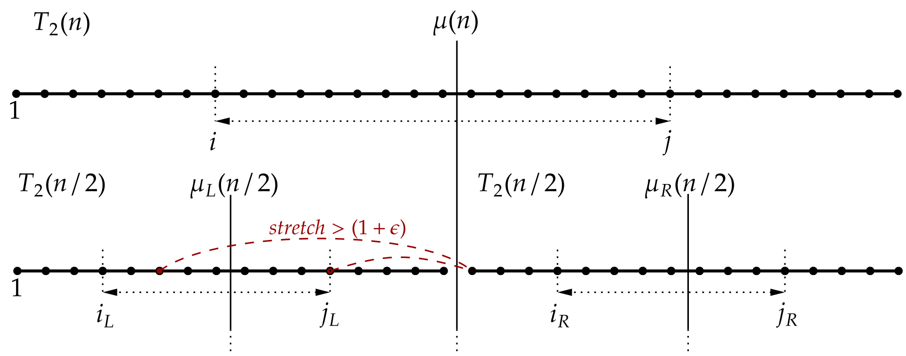

Figure 1 An illustration of the first two levels of the recurrence for the lower bound for $k = 2$ and $\epsilon = 1/2$. We split the interval $U(1, n)$ into two disjoint parts. In Lemma 10, we show that there will be at least $\Omega(n)$ cross edges, which are the spanner edges having endpoints in both parts. The values i_L and j_L are set according to Corollary 8 so that the spanner edges crossing $\mu(n)$ cannot be used for the left set; otherwise the resulting stretch will be bigger than $1 + \epsilon$.

▶ **Lemma 11** (Proof omitted; see the full version [26]). *For any positive integer n, and any two integers l, r such that $r = l + n - 1$, let $U(l, r)$ be a uniform line metric with n points and let X be a subspace of $U(l, r)$ which contains at least $\frac{31}{32}n$ points. Then, for any choice of $\epsilon \in [0, 1/2]$, any spanner on X with hop-diameter 2 and stretch $1 + \epsilon$ contains at least $T_2'(n) \geq 0.48 \cdot \frac{n}{16} \log n$ edges which have both endpoints inside $[l, r]$.*

Completing the proof of Theorem 9. Note that $\alpha_2(n) = \lceil \log n \rceil$ and hence, we will show that $T'_2(n) \geq \frac{n}{256}\lceil \log n \rceil$. Suppose without loss of generality that we are working on any t-sparse line metric with n points, $U((1,N),t)$, where $N = nt$. Let H be an arbitrary $(1+\epsilon)$-spanner for $U((1,N),t)$ with $((1,N),t)$-global hop-diameter 2. We would like to lower bound the number of $((l,r),t)$-global edges required for H.

Since $\epsilon \in [0,1/2]$, every two consecutive points in $U((1,N),t)$, except for the leftmost and the rightmost two, have to be connected by a spanner path which has all its endpoints inside the interval $[1,N]$. This implies that the number of spanner edges is at least $n-3$, which is in turn greater than $(n/16)\log n$, for any $64 \leq n \leq 127$.

Let $M = \lfloor n/2 \rfloor t$ and let L be the set of $((l,r),t)$-intervals that are fully inside $[N/4, M]$ and R be the set of $((l,r),t)$-intervals that are fully inside $[M, 3N/4]$. In that case, the number of $((l,r),t)$-intervals inside L can be lower bounded by $|L| \geq \lfloor (M - N/4 + 1)/t \rfloor \geq n/4 - 2$, which is the bound that we used for L. Similarly, we obtain that $|R| \geq n/4 - 1$. The cross edges will be those edges that contain one endpoint in $[1,M]$ and the other endpoint in $[M+1,N]$. It follows that the cross edges are also $((l,r),t)$-global edges. The same argument can be applied to lower bound the number of cross edges, implying the lower bound on the number of $((l,r),t)$-global edges. The same proof as in Lemma 11 gives $T'_2(n) \geq 0.48 \cdot \frac{n}{16}\log n \geq \frac{n}{256}\lceil \log n \rceil$, when $n \geq 1000$, as desired. ◀

3.2 Hop diameter 3

▶ **Theorem 12.** *For any two positive integers $n \geq 1000$ and t, and any two integers l, r such that $r = l + nt - 1$, let $U((l,r),t)$ be a t-sparse line metric with n points and let X be a subspace of $U((l,r),t)$ which contains at least $\frac{127}{128}n$ points. Then, for any choice of $\epsilon \in [0,1/2]$, any spanner on X with $((l,r),t)$-global hop-diameter 3 and stretch $1+\epsilon$ contains at least $T'_3(n) \geq \frac{n}{1024} \cdot \alpha_3(n)$ $((l,r),t)$-global edges which have both endpoints inside $[l,r]$.*

The theorem is proved in three steps. First, we prove Lemma 13, which concerns uniform line metrics. Then, we prove Lemma 16 for a subspace that contains at least $31/32$ fraction of the points of the original metric. In the third step, we observe that the same argument applies for t-sparse line metrics.

▶ **Lemma 13.** *For any positive integer n, and any two integers l, r such that $r = l + n - 1$, let $U(l,r)$ be a uniform line metric with n points. Then, for any choice of $\epsilon \in [0,1/2]$, any spanner on $U(l,r)$ with hop-diameter 3 and stretch $1+\epsilon$ contains at least $T_3(n) \geq \frac{n}{40}\log\log n$ edges which have both endpoints inside $[l,r]$.*

Proof. Suppose without loss of generality that we are working on the uniform line metric $U(1,n)$. Let H be an arbitrary $(1+\epsilon)$-spanner for $U(1,n)$ with hop-diameter 3.

For the base case, we assume that $11 \leq n \leq 127$. We have that $\frac{n}{40}\log\log n < n-1$, which is a trivial lower bound on the number of edges of H, since every two consecutive points have to be connected via a direct edge.

We now assume that $n \geq 128$. Divide the the interval $[1,n]$ into consecutive subintervals containing $b := \lfloor \sqrt{n} \rfloor$ points: $[1,b], [b+1,2b]$, etc. Our goal is to show that the number of spanner edges is lower bounded by $T_3(n)$, which satisfies recurrence $T_3(n) = \left\lfloor \frac{n}{\lfloor \sqrt{n} \rfloor} \right\rfloor \cdot T_3\left(\lfloor \sqrt{n} \rfloor\right) + n/18$, with the base case $T_3(n) = (n/40)\log\log n$ when $n < 128$.

For any j such that $1 \leq j \leq \lfloor n/b \rfloor$, the interval spanned by the jth subinterval is $[(j-1)b+1, jb]$. Using the induction hypothesis, any spanner on $U((j-1)b+1, jb)$ contains at least $T_3(b)$ edges that are inside $[(j-1)b+1, jb]$. This means that all the subintervals will contribute at least $\lfloor n/b \rfloor \cdot T_3(b)$ spanner edges that are mutually disjoint and in addition do not go outside of $[1, n]$. We will show that there are $\Omega(n)$ edges that have endpoints in two different subintervals, called *cross edges*. By definition, the set of cross edges is disjoint from the set of spanner edges considered in the term $\lfloor n/b \rfloor \cdot T_3(b)$.

Consider the points that are within interval $[n/4, 3n/4]$. From Corollary 8, since n is sufficiently large, we know that any $(1+\epsilon)$-spanner path connecting two points in $[n/4, 3n/4]$ has to have all its points inside $[1, n]$.

We call a point *global* if it is adjacent to at least one cross edge. Otherwise, the point is *non-global*. The following two claims bound the number of cross edges induced by global and non-global points, respectively.

▷ Claim 14. Suppose that among points inside interval $[n/4, 3n/4]$, m of them are global. Then, they induce at least $m/2$ spanner edges.

The claim is true since each global point contributes at least one cross edge and each edge is counted at most twice.

▷ Claim 15. Suppose that among points inside interval $[n/4, 3n/4]$, m of them are non-global. Then, they induce at least $\binom{m/\sqrt{n}}{2}$ cross edges.

Proof. Consider two sets A and B such that A contains a non-global point $a \in [n/4, 3n/4]$ and B contains a non-global point $b \in [n/4, 3n/4]$. Since a is non-global, it can be connected via an edge either to a point inside of A or to a point outside of $[1, n]$. Similarly, b can be connected to either a point inside of B or to a point outside of $[1, n]$. From Corollary 8, and since $a, b \in [n/4, 3n/4]$, we know that every spanner path with stretch $(1+\epsilon)$ connecting a and b has to use points inside $[1, n]$. This means that the spanner path with stretch $(1+\epsilon)$ has to have a form (a, a', b', b), where $a' \in A$ and $b' \in B$. In other words, we have to connect points a' and b' using a cross edge; furthermore every pair of intervals containing at least one non-global point induce one such edge and for every pair this edge is different.

Each interval contains at most $b = \lfloor \sqrt{n} \rfloor$ non-global points, so the number of sets containing at least one non-global point is at least m/b. Interconnecting all the sets requires $\binom{m/b}{2} \geq \binom{m/\sqrt{n}}{2}$ edges. ◁

The number of points inside $[n/4, 3n/4]$ is at least $n/2 + 1$, but we shall use a slightly weaker lower bound of $15n/32$. We consider two complementary cases. In the first case, at least $1/4$ of $15n/32$ points are global. Claim 14 implies that the number of the cross edges induced by these points is at least $15n/256$. The other case is that at least $3/4$ fraction of $15n/32$ points are non-global. Claim 15 implies that for a sufficiently large n, the number of cross edges induced by these points can be lower bounded by $15n/256$ as well. In other words, we have shown that in both cases, the number of cross edges is at least $\frac{15}{256}n > \frac{n}{18}$. Thus, we have: $T_3(n) \geq \left\lfloor \frac{n}{\lfloor \sqrt{n} \rfloor} \right\rfloor \cdot T_3\left(\lfloor \sqrt{n} \rfloor\right) + \frac{n}{18} \geq \lfloor \sqrt{n} \rfloor \cdot \frac{\lfloor \sqrt{n} \rfloor}{40}(\log\log \lfloor \sqrt{n} \rfloor) + \frac{n}{18}$, which is at most $\frac{n}{40} \log\log n$, as claimed. ◀

▶ **Lemma 16** (Proof omitted; see the full version [26]). *For any positive integer n, and any two integers l, r such that $r = l + n - 1$, let $U(l, r)$ be a uniform line metric with n points and let X be a subspace of $U(l, r)$ which contains at least $\frac{127}{128}n$ points. Then, for any choice of $\epsilon \in [0, 1/2]$, any spanner on X with hop-diameter 3 and stretch $1+\epsilon$ contains at least $T_3'(n) \geq 0.18 \cdot \frac{n}{40} \log\log n$ edges which have both endpoints inside $[l, r]$.*

Completing the proof of Theorem 12. Note that $\alpha_3(n) = \lceil \log\log n \rceil$ and hence, we will show that $T_3'(n) \geq \frac{n}{1024} \cdot \lceil \log\log n \rceil$. Suppose without loss of generality that we are working on any t-sparse line metric with n points, $U((1,N),t)$, where $N = nt$. Let H be an arbitrary $(1+\epsilon)$-spanner for $U((1,N),t)$ with $((1,N),t)$-global hop-diameter 3. We would like to lower bound the number of $((l,r),t)$-global edges required for H.

Since $\epsilon \in [0,1/2]$, every two consecutive points in $U((1,N),t)$, except for the leftmost and the rightmost two, have to be connected by a spanner path which has all its endpoints inside the interval $[1,N]$. This implies that the number of spanner edges is at least $n-3$, which is in turn greater than $(n/40)\log\log n$, for any $11 \leq n \leq 127$.

Let consider the set of $((l,r),t)$-intervals that are fully inside $[N/4, 3N/4]$. The number of such intervals can be lower bounded by $((3N/4 - N/4)/t - 2 \geq n/2 - 2$, which is larger than the bound of $15n/32$, which we used. The cross edges will become $((1,N),t)$-global edges and the same argument can be applied to lower bound their number. The same proof in Lemma 16 gives:

$$T_3'(n) \geq 0.18 \cdot \frac{n}{40} \log\log n \geq \frac{n}{1024} \cdot \lceil \log\log n \rceil$$

when $n \geq 1000$, as desired. ◀

4 Lower bound for constant hop-diameter

We proceed to prove our main result, which is a generalization of Theorem 1. In particular, invoking Theorem 17 stated below where X is the uniform line metric $U(1,n)$ gives Theorem 1.

▶ **Theorem 17.** *For any two positive integers $n \geq 1000$ and t, and any two integers l, r such that $r = l + nt - 1$, let $U((l,r),t)$ be a t-sparse line metric with n points and let X be a subspace of $U((l,r),t)$ which contains at least $n(1 - \frac{1}{2^{k+4}})$ points. Then, for any choice of $\epsilon \in [0,1/2]$ and any integer $k \geq 2$, any spanner on X with $((l,r),t)$-global hop-diameter k and stretch $1+\epsilon$ contains at least $T_k'(n) \geq \frac{n}{2^{6\lfloor k/2 \rfloor + 4}} \cdot \alpha_k(n)$ $((l,r),t)$-global edges which have both endpoints inside $[l,r]$.*

Proof. We will prove the theorem by double induction on $k \geq 2$ and n. The base case for $k = 2$ and $k = 3$ and every n is proved in Theorems 9 and 12, respectively.

For every $k \geq 4$, we shall prove the following two assertions.

1. For any two positive integers n and t, and any two integers l, r such that $r = l + nt - 1$, let $U((l,r),t)$ be a t-sparse line metric with n points. Then, for any choice of $\epsilon \in [0,1/2]$, any spanner on $U((l,r),t)$ with $((l,r),t)$-global hop-diameter k and stretch $1+\epsilon$ contains at least $T_k(n) \geq \frac{n}{2^{6\lfloor k/2 \rfloor + 2}} \alpha_k(n)$ $((l,r),t)$-global edges which have both endpoints inside $[l,r]$.
2. For any two positive integers n and t, and any two integers l, r such that $r = l + nt - 1$, let $U((l,r),t)$ be a t-sparse line metric with n points and let X be a subspace of $U((l,r),t)$ which contains at least $n(1 - \frac{1}{2^{k+4}})$ points. Then, for any choice of $\epsilon \in [0,1/2]$, any spanner on X with $((l,r),t)$-global hop-diameter k and stretch $1+\epsilon$ contains at least $T_k'(n) \geq \frac{n}{2^{6\lfloor k/2 \rfloor + 4}} \cdot \alpha_k(n)$ $((l,r),t)$-global edges which have both endpoints inside $[l,r]$.

For every $k \geq 4$, we first prove the first assertion, which relies on the second assertion for $k-2$. Then, we prove the second assertion which relies on the first assertion for k. We proceed to prove assertion 1.

Proof of assertion 1. Suppose without loss of generality that we are working on any t-sparse line metric $U((1,N),t)$. Let H be an arbitrary $(1+\epsilon)$-spanner for $U((1,N),t)$ with $((1,N),t)$-global hop-diameter k.

Let M be $A((k-2)/2,4)$ if k is even and $B(\lfloor(k-2)/2\rfloor,4)$ if k is odd. For the base case take $4 \le n < \max(M,10000)$. We consider $n-2$ points in $U((1,N),t)$: all the points from the metric, excluding the leftmost and the rightmost one. Since $\epsilon \in [0,1/2]$, every two consecutive points among the considered $n-2$ points have to be connected by a spanner path which has all its endpoints inside the interval $[1,N]$. This implies that the number of spanner edges is at least $n-3$. Then $\frac{n}{2^{6\lfloor k/2\rfloor+2}}\alpha_k(n)$, which is at most $\frac{n}{2^{6\lfloor k/2\rfloor+2}}\log^*(n) \le n-3$.

Next, we prove the induction step. We shall assume the correctness of the two statements: (i) for k and all smaller values of n, and (ii) for $k' < k$ and all values of n. Let $N := nt$ and let $b := \alpha_{k-2}(n)$. Divide the the interval $[1,N]$ into consecutive $((1,N),bt)$-intervals containing b points: $[1,bt],[bt+1,2bt]$, etc. We would like to prove that the number of spanner edges is lower bounded by recurrence

$$T_k(n) = \left\lfloor \frac{n}{\alpha_{k-2}(n)} \right\rfloor \cdot T_k(\alpha_{k-2}(n)) + \frac{n}{2^{6\lfloor k/2\rfloor+1}},$$

with the base case $T_k(n) = \frac{n}{2^{6\lfloor k/2\rfloor+2}}\alpha_k(n)$ for $n \le 10000$.

There are $\lfloor n/b\rfloor$ $((1,N),bt)$-intervals containing exactly b points. For any j such that $1 \le j \le \lfloor n/b\rfloor$, the jth $((1,N),bt)$-interval is $[(j-1)bt+1, jbt]$. Using inductively the assertion 1 for k and a value $b < n$, any spanner on $U((j-1)bt+1, jbt)$ contains at least $T_k(b)$ edges that are inside $[(j-1)bt+1, jbt]$. This means that all the $((1,N),bt)$-intervals will contribute at least $\lfloor n/b\rfloor \cdot T_k(b)$ spanner edges that are mutually disjoint and in addition do not go outside of $[1,N]$.

We will show that there are $\Omega(n/2^{3k})$ edges that have endpoints in two different $((1,N),bt)$-intervals, i.e. edges that are $((1,N),bt)$-global. Since these edges are $((1,N),bt)$-global, they are disjoint from the spanner edges considered in the term $\lfloor n/b\rfloor \cdot T_3(b)$. We shall focus on points that are inside $((1,N),bt)$-intervals fully inside $[N/4,3N/4]$; denote the number of such points by p. We have $p \ge n/2 - 2\alpha_{k-2}(n)$, but we will use a weaker bound:

$$p \ge n/4. \tag{1}$$

▶ **Definition 18.** *A point that is incident on at least one* $((1,N),bt)$*-global edge is called a* $((1,N),bt)$*-global point.*

Among the p points inside inside $[N/4,3N/4]$, denote by p' the number of $((1,N),bt)$-global points. Let $p'' = p - p'$, and m be the number of $((1,N),bt)$-global edges incident on the p points. Since each $((1,N),bt)$-global point contributes at least one $((1,N),bt)$-global edge and each such edge is counted at most twice, we have

$$m \;\ge\; p'/2. \tag{2}$$

Next, we prove that

$$m \;\ge\; \frac{n}{2^{6\lfloor k/2\rfloor+1}}, \quad \text{if } \left\lceil \frac{p''}{b} \right\rceil \;\ge\; \left(1 - \frac{1}{2^{k+2}}\right) \cdot \left\lceil \frac{p}{b} \right\rceil \tag{3}$$

Recall that we have divided $[1,N]$ into consecutive $((1,N),bt)$-intervals containing $b := \alpha_{k-2}(n)$ points. Consider now all the $((1,N),bt)$-intervals that are fully inside $[N/4,3N/4]$, and denote this collection of $((1,N),bt)$-intervals by $\mathcal{C}$. Let l' (resp. r') be the leftmost (resp. rightmost) point of the leftmost (resp. rightmost) interval in $\mathcal{C}$; note that l' and r' may not coincide with points of the input metric, they are simply the leftmost and rightmost boundaries of the intervals in $\mathcal{C}$.

Constructing a new line metric. For each $((1,N),bt)$-interval I in $\mathcal{C}$, if I contains a point that is not $((1,N),bt)$-global, assign an arbitrary such point in I as its representative; otherwise, assign an arbitrary point as its representative. The collection $\mathcal{C}$ of $((1,N),bt)$-intervals, together with the set of representatives uniquely defines (bt)-sparse line metric, $U((l',r'),bt)$. This metric has $\lceil p/b \rceil$ $((1,N),bt)$-intervals, since there are $\lceil p/b \rceil$ intervals covering p points in the input t-sparse metric $U((1,N),t)$ inside the interval $[N/4,3N/4]$. Recall from Definition 4 that a bt-sparse metric is uniquely defined given its $((1,N),bt)$-intervals and representatives. Let X be the subspace of $U((l',r'),bt)$ induced by the representatives of all intervals in $\mathcal{C}$ that contain points that are not $((1,N),bt)$-global and using Equation (3), we have

$$|X| \;\geq\; \left\lceil \frac{p''}{b} \right\rceil \;\geq\; \left(1-\frac{1}{2^{k+2}}\right)\cdot \left\lceil \frac{p}{b} \right\rceil \tag{4}$$

Recall that H is an arbitrary $(1+\epsilon)$-spanner for $U((1,N),t)$ with $((1,N),t)$-global hop-diameter k. Let a and b be two arbitrary points in X, and denote their corresponding $((1,N),bt)$-intervals by A and B, respectively. Since a (reps., b) is not $((1,N),bt)$-global, it can be adjacent either to points outside of $[1,N]$ or to points inside A (resp., B). By Corollary 8 and since $a,b \in [N/4,3N/4]$, any spanner path with stretch $(1+\epsilon)$ connecting a and b must remain inside $[1,N]$. Hence, any $(1+\epsilon)$-spanner path in H between a and b is of the form $(a,a',\ldots,b',b)$, where $a' \in A$ (resp. $b' \in B$). Consider now the same path in the metric X. It has at most k hops, where the first and the last edges are not $((1,N),bt)$-global. Thus, although this path contains at most k $((1,N),t)$-global edges in $U((1,N),t)$, it has at most $k-2$ $((1,N),bt)$-global edges in X. It follows that H is a (Steiner) $(1+\epsilon)$-spanner with $((1,N),bt)$-global hop-diameter $k-2$ for X. See Figure 2 for an illustration.

Denote by $n' := \lceil p/b \rceil$ the number of points in $U((l',r'),bt)$. Since $p \geq n/4$, it follows that $n' \geq \lceil n/(4b) \rceil$. By (4), X is a subspace of $U((l',r'),bt)$, and its size is at least a $(1-1/2^{k+2})$-fraction (i.e., a $(1-1/2^{(k-2)+4})$-fraction) of that of $U((l',r'),bt)$. Hence, by the induction hypothesis of assertion 2 for $k-2$, we know that any spanner on X with $((l',r'),bt)$-global hop-diameter $k-2$ and stretch $1+\epsilon$ contains at least $T'_{k-2}(n') \geq \frac{n'}{2^{6\lfloor (k-2)/2 \rfloor+4}} \cdot \alpha_{k-2}(n')$ $((l',r'),bt)$-global edges which have both endpoints inside $[l',r']$. Since every $((l',r'),bt)$-global edge is also a $((1,N),bt)$-global edge, we conclude with the following lower bound on the number of $((1,N),bt)$-global edges required by H:

$$\begin{aligned}
T'_{k-2}\left(n'\right) &\geq \frac{n'}{2^{6\lfloor (k-2)/2 \rfloor+4}} \cdot \alpha_{k-2}\left(n'\right)\\
&\geq \frac{n}{4\cdot 2^{6\lfloor (k-2)/2 \rfloor+4}\cdot \alpha_{k-2}(n)} \cdot \alpha_{k-2}\left(\left\lceil \frac{n}{4\alpha_{k-2}(n)} \right\rceil\right)\\
&\geq \frac{n}{8\cdot 2^{6\lfloor (k-2)/2 \rfloor+4}}\\
&= \frac{n}{2^{6\lfloor k/2 \rfloor+1}}
\end{aligned}$$

The last inequality follows since, when $k \geq 4$, the ratio between $\alpha_{k-2}(\lceil n/4\alpha_{k-2}(n) \rceil)$ and $\alpha_{k-2}(n)$ can be bounded by $1/2$ for sufficiently large n (i.e. larger than the value considered in the base case). In other word, we have shown that whenever $\lceil p''/b \rceil \;\geq\; (1-1/2^{k+2})\cdot\lceil p/b \rceil$, the number of the $((1,N),bt)$-global edges incident on the p points inside $[N/4,3N/4]$ is lower bounded by $n/2^{6\lfloor k/2 \rfloor+1}$; we have thus proved (3).

Recall (see (1)) that we lower bounded the number p of points inside $[N/4,3N/4]$ as $p \geq n/4$. We consider two complementary cases: either $\lceil p''/b \rceil \;\geq\; (1-1/2^{k+2})\cdot\lceil p/b \rceil$, or $\lceil p''/b \rceil \;<\; (1-1/2^{k+2})\cdot\lceil p/b \rceil$, where p'' is the number of points in $[N/4,3N/4]$ that are not $((1,N),bt)$-global. In the former case (i.e. when $\lceil p''/b \rceil \;\geq\; (1-1/2^{k+2})$), by (3), we have the number of $((1,N),bt)$-global edges is lower bounded by $n/2^{6\lfloor k/2\rfloor+1}$. In the latter case, we have $\frac{p-p'}{b} - 1 < \left\lfloor \frac{p-p'}{b} \right\rfloor = \frac{p''}{b} < \left(1 - \frac{1}{2^{k+2}}\right) \cdot \lceil \frac{p}{b} \rceil < \left(1 - \frac{1}{2^{k+2}}\right) \cdot \frac{p}{b} + 1$. In other words, we can lower bound p' by $p/2^{k+2} - 2b$. From (2) and using that $p \geq n/4$, the number of $((1,N),bt)$-global edges is lower bounded by $n/2^{k+5} - \alpha_{k-2}(n)$. Since the former bound is always smaller for n sufficiently large (i.e. larger than the value considered in the base case), we shall use it as a lower bound on the number of $((1,N),bt)$-global edges required by H. We note that every $((1,N),bt)$-global edge is also $((1,N),t)$-global, as required by assertion 1. It follows that

$$\begin{aligned} T_k(n) &\geq \left\lfloor \frac{n}{\alpha_{k-2}(n)} \right\rfloor \cdot \frac{\alpha_{k-2}(n)}{2^{6\lfloor k/2\rfloor+2}} \cdot \alpha_k(\alpha_{k-2}(n)) + \frac{n}{2^{6\lfloor k/2\rfloor+1}} \\ &\geq \left(\frac{n}{\alpha_{k-2}(n)} - 1\right) \cdot \frac{\alpha_{k-2}(n)}{2^{6\lfloor k/2\rfloor+2}} \cdot (\alpha_k(n) - 1) + \frac{n}{2^{6\lfloor k/2\rfloor+1}} \\ &\geq \frac{n}{2^{6\lfloor k/2\rfloor+2}} \alpha_k(n) \end{aligned}$$

For the second inequality we have used that $\alpha_k(n) = 1 + \alpha_k(\alpha_{k-2}(n))$, and for the third, the fact that $\alpha_{k-2}(n) \cdot (\alpha_k(n) - 1) \leq n$ for sufficiently large n (i.e. larger than the value considered in the base case). This concludes the proof of assertion 1.

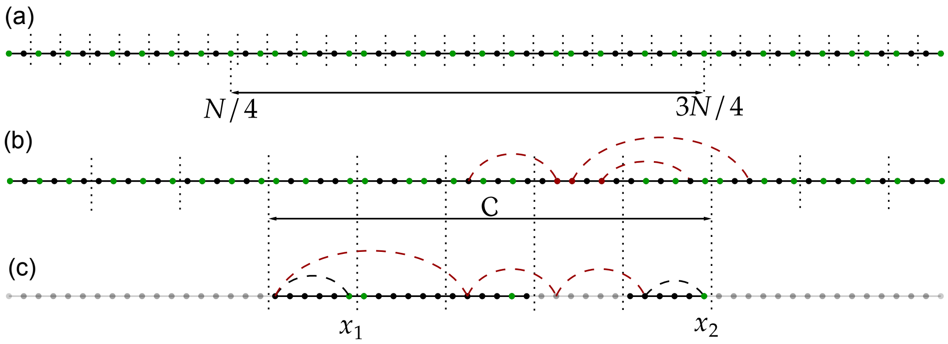

Figure 2 Constructing a new line metric and invoking the induction hypothesis. **(a)** We have $n = 32$, $k = 5$, and a 2-sparse line metric $U((1,64),2)$ with representatives of each $((1,64),2)$-interval highlighted in green. **(b)** Since $b = \alpha_{k-2}(n) = 3$, we consider a collection of $((1,64),6)$-global intervals inside $[N/4, 3N/4]$, denoted by $\mathcal{C}$. The seventh block contains only $((1,64),6)$-global points (highlighted in red) as each of them is incident on a $((1,64),6)$-global edge. **(c)** The new line metric is 6-sparse line metric $U((19,48),6)$ consisting of 4 green points. Finally, we use the induction hypothesis of assertion 2 for $k = 3$ to lower bound the number of $((1,N),6)$-global edges. A spanner path between x_1 and x_2 consisting of 5 edges, 3 of which are $((1,N),6)$ global is depicted.

Proof of assertion 2. Suppose without loss of generality that we are working on any t-sparse line metric $U((1,N),t)$. Let H be an arbitrary $(1+\epsilon)$-spanner for $U(1,N)$ with $((1,N),t)$-global hop-diameter k. We shall inductively assume the correctness of assertion 1 and assertion 2: (i) for k and all smaller values of n, and (ii) for $k' < k$ and all values of n.

Recall the recurrence we used in the proof of assertion 1, $T_k(n) = \lfloor n/\alpha_{k-2}(n) \rfloor \cdot T_k(\alpha_{k-2}(n)) + \frac{n}{2^{6\lfloor k/2\rfloor+1}}$, which provides a lower bound on the number of $((l,r),t)$-global edges of H. The base case for this recurrence is whenever $n < 10000$. Consider the recursion tree of $T_k(n)$ and denote its depth by ℓ and the number of nodes at depth i by c_i. In addition, denote by $n_{i,j}$ the number of points in the jth interval of the ith level and by $e_{i,j}$ the number of $((1,N),t)$-global edges contributed by this interval. We have that the contribution of an interval is $n_{i,j}/2^{6\lfloor k/2\rfloor+1}$. By definition, we have $T_k(n) = \sum_{i=1}^{\ell} \sum_{j=1}^{c_i} e_{i,j} \geq \frac{n}{2^{6\lfloor k/2\rfloor+2}} \alpha_k(n)$.

Let H' be any $(1+\epsilon)$ spanner on X with $((1,N),t)$-global hop-diameter k. To lower bound the number of spanner edges in H', we now consider the same recursion tree, but take into consideration the fact that we are working on metric X, which is a subspace of $U((1,N),t)$. This means that at each level of recursion, instead of n points, there is at least $n(1-1/2^{k+4})$ points in X. The contribution of the jth interval in the ith level is denoted by $e'_{i,j}$. We call the jth interval in the ith level *good* if it contains at least $n_{i,j}(1-1/2^{k+3})$ points from X. (Recall that we have used $n_{i,j}$ to denote the number of points from $U(l,r)$ in the jth interval of the ith level.) From the definition of good interval and the fact that each level of recurrence contains at least $n(1-1/2^{k+4})$ points, it follows that there are at least $n/2$ points contained in the good intervals at the ith level. Denote the collection of all the good intervals at the ith level by Γ_i.

Recall that we are working with recurrence $T_k(n) = \lfloor n/\alpha_{k-2}(n)\rfloor \cdot T_k(\alpha_{k-2}(n)) + \frac{n}{2^{6\lfloor k/2\rfloor+1}}$. In particular, in the first level of recurrence, we consider the contribution of n points, whereas in the second level, we consider the contribution of $\lfloor n/\alpha_{k-2}(n)\rfloor \cdot \alpha_{k-2}(n)$ points. Denote by n_i the number of points whose contribution we consider in the ith level of recurrence. Then, we have $n_1 = n$, $n_2 = \lfloor n/\alpha_{k-2}(n)\rfloor \cdot \alpha_{k-2}(n) \geq n - \alpha_{k-2}(n)$. Denote by $\alpha^{(j)}_{k-2}(n)$ value of $\alpha_{k-2}(\cdot)$ iterated on n, i.e. $\alpha^{(0)}_{k-2}(n) = n$, $\alpha^{(1)}_{k-2}(n) = \alpha_{k-2}(n)$, $\alpha^{(2)}_{k-2}(n) = \alpha_{k-2}(\alpha_{k-2}(n))$, etc. In general, for $i \geq 2$, we have $n_i \geq n - \sum_{j=2}^{i} \frac{n\alpha^{(j-1)}_{k-2}(n)}{\alpha^{(j-2)}_{k-2}(n)} \geq n - n\cdot\sum_{j=2}^{i} \frac{\lceil \log^{(j-1)}(n)\rceil}{\lceil \log^{(j-2)}(n)\rceil}$. We observe that there is an exponential decay between the numerator and denominator of terms in each summand and that terms grow with j. Since we do not consider intervals in the base case, we also know that $\lceil \log^{(i-1)}(n)\rceil \geq 10000$, meaning that the largest term in the sum is $10000/2^{9999}$. By observing that every two consecutive terms increase by a factor larger than 2, we conclude that $n_i \geq 0.99n$. Since at each level there are at least $n/2$ points inside of good intervals, this means that there are at least $0.49n$ points inside of good intervals which were not ignored. Denote by Γ_i the set of good intervals in the ith level whose contribution is not ignored. Then we have $T'_k(n) = \sum_{i=1}^{\ell}\sum_{j=1}^{c'_i} e'_{i,j} \;\geq\; \sum_{i=1}^{\ell}\sum_{j\in\Gamma_i} e_{i,j} \;\geq\; 0.49\cdot T_k(n) \geq \frac{n}{2^{6\lfloor k/2\rfloor+4}}\alpha_k(n)$. This concludes the proof of assertion 2. We have thus completed the inductive step for k. ◀

References

1 Ittai Abraham and Dahlia Malkhi. Compact routing on euclidian metrics. In *PODC*, pages 141–149. ACM, 2004.

2 Pankaj K. Agarwal, Yusu Wang, and Peng Yin. Lower bound for sparse euclidean spanners. In *SODA*, pages 670–671. SIAM, 2005.

3 Noga Alon and Baruch Schieber. Optimal preprocessing for answering on-line product queries. Technical report, Tel Aviv University, 1987.

4 Ingo Althöfer, Gautam Das, David P. Dobkin, Deborah Joseph, and José Soares. On sparse spanners of weighted graphs. *Discret. Comput. Geom.*, 9:81–100, 1993.

5 Sunil Arya, Gautam Das, David M. Mount, Jeffrey S. Salowe, and Michiel H. M. Smid. Euclidean spanners: short, thin, and lanky. In *STOC*, pages 489–498. ACM, 1995.

6 Sunil Arya, David M. Mount, and Michiel H. M. Smid. Randomized and deterministic algorithms for geometric spanners of small diameter. In *FOCS*, pages 703–712. IEEE Computer Society, 1994.

7 Yair Bartal, Nova Fandina, and Ofer Neiman. Covering metric spaces by few trees. In *ICALP*, volume 132 of *LIPIcs*, pages 20:1–20:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.

8 Arnab Bhattacharyya, Elena Grigorescu, Kyomin Jung, Sofya Raskhodnikova, and David P. Woodruff. Transitive-closure spanners. In *SODA*, pages 932–941. SIAM, 2009.

9 Hans L. Bodlaender, Gerard Tel, and Nicola Santoro. Trade-offs in non-reversing diameter. *Nord. J. Comput.*, 1(1):111–134, 1994.

10 Paul B. Callahan and S. Rao Kosaraju. Faster algorithms for some geometric graph problems in higher dimensions. In *SODA*, pages 291–300. ACM/SIAM, 1993.

11 T.-H. Hubert Chan and Anupam Gupta. Small hop-diameter sparse spanners for doubling metrics. *Discret. Comput. Geom.*, 41(1):28–44, 2009.

12 Bernard Chazelle. Computing on a free tree via complexity-preserving mappings. *Algorithmica*, 2:337–361, 1987.

13 Bernard Chazelle and Burton Rosenberg. The complexity of computing partial sums off-line. *Int. J. Comput. Geom. Appl.*, 1(1):33–45, 1991.

14 Danny Z. Chen, Gautam Das, and Michiel H. M. Smid. Lower bounds for computing geometric spanners and approximate shortest paths. *Discret. Appl. Math.*, 110(2-3):151–167, 2001.

15 Paul Chew. There is a planar graph almost as good as the complete graph. In *SCG*, pages 169–177. ACM, 1986.

16 Gautam Das and Giri Narasimhan. A fast algorithm for constructing sparse euclidean spanners. *Int. J. Comput. Geom. Appl.*, 7(4):297–315, 1997.

17 Gautam Das, Giri Narasimhan, and Jeffrey S. Salowe. A new way to weigh malnourished euclidean graphs. In *SODA*, pages 215–222. ACM/SIAM, 1995.

18 Yefim Dinitz, Michael Elkin, and Shay Solomon. Low-light trees, and tight lower bounds for euclidean spanners. *Discret. Comput. Geom.*, 43(4):736–783, 2010.

19 Michael Elkin and Shay Solomon. Optimal euclidean spanners: Really short, thin, and lanky. *J. ACM*, 62(5):35:1–35:45, 2015.

20 Joachim Gudmundsson, Christos Levcopoulos, Giri Narasimhan, and Michiel H. M. Smid. Approximate distance oracles for geometric graphs. In *SODA*, pages 828–837. ACM/SIAM, 2002.

21 Joachim Gudmundsson, Christos Levcopoulos, Giri Narasimhan, and Michiel H. M. Smid. Approximate distance oracles for geometric spanners. *ACM Trans. Algorithms*, 4(1):10:1–10:34, 2008.

22 Joachim Gudmundsson, Giri Narasimhan, and Michiel H. M. Smid. Fast pruning of geometric spanners. In *STACS*, volume 3404 of *Lecture Notes in Computer Science*, pages 508–520. Springer, 2005.

23 Yehuda Hassin and David Peleg. Sparse communication networks and efficient routing in the plane. *Distributed Comput.*, 14(4):205–215, 2001.

24 Omri Kahalon, Hung Le, Lazar Milenkovic, and Shay Solomon. Can't see the forest for the trees: Navigating metric spaces by bounded hop-diameter spanners. *CoRR*, abs/2107.14221, 2021. arXiv:2107.14221.

25 J. Mark Keil and Carl A. Gutwin. Classes of graphs which approximate the complete euclidean graph. *Discret. Comput. Geom.*, 7:13–28, 1992.

26 Hung Le, Lazar Milenkovic, and Shay Solomon. Sparse euclidean spanners with tiny diameter: A tight lower bound. *CoRR*, abs/2112.09124, 2021. arXiv:2112.09124.

27 Hung Le and Shay Solomon. Truly optimal euclidean spanners. In *FOCS*, pages 1078–1100. IEEE Computer Society, 2019.

28 Christos Levcopoulos, Giri Narasimhan, and Michiel H. M. Smid. Efficient algorithms for constructing fault-tolerant geometric spanners. In *STOC*, pages 186–195. ACM, 1998.

29 Yishay Mansour and David Peleg. An approximation algorithm for minimum-cost network design. In *Robust Communication Networks: Interconnection and Survivability*, volume 53 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 97–106. DIMACS/AMS, 1998.

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

31 Mihai Patrascu and Erik D. Demaine. Tight bounds for the partial-sums problem. In *SODA*, pages 20–29. SIAM, 2004.

32 Satish Rao and Warren D. Smith. Approximating geometrical graphs via "spanners" and "banyans". In *STOC*, pages 540–550. ACM, 1998.

33 Shay Solomon. Sparse euclidean spanners with tiny diameter. *ACM Trans. Algorithms*, 9(3):28:1–28:33, 2013.

34 Shay Solomon. From hierarchical partitions to hierarchical covers: optimal fault-tolerant spanners for doubling metrics. In *STOC*, pages 363–372. ACM, 2014.

35 Shay Solomon and Michael Elkin. Balancing degree, diameter and weight in euclidean spanners. In *ESA (1)*, volume 6346 of *Lecture Notes in Computer Science*, pages 48–59. Springer, 2010.

36 Robert Endre Tarjan. Applications of path compression on balanced trees. *J. ACM*, 26(4):690–715, 1979.

37 Mikkel Thorup. On shortcutting digraphs. In *WG*, volume 657 of *Lecture Notes in Computer Science*, pages 205–211. Springer, 1992.

38 Mikkel Thorup. Shortcutting planar digraphs. *Comb. Probab. Comput.*, 4:287–315, 1995.

39 Andrew Chi-Chih Yao. Space-time tradeoff for answering range queries (extended abstract). In *STOC*, pages 128–136. ACM, 1982.