

Sparse Euclidean Spanners with Tiny Diameter: A Tight Lower Bound

Hung Le¹, Lazar Milenković², and Shay Solomon³

¹University of Massachusetts Amherst

^{2,3}Tel Aviv University

May 3, 2023

Abstract

In STOC'95 [ADM⁺95] Arya et al. showed that any set of n points in $\mathbb{R}^d$ 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 $\alpha_0(n) = \lceil n/2 \rceil$, $\alpha_1(n) = \lceil \sqrt{n} \rceil$, $\alpha_2(n) = \lceil \log n \rceil$, $\alpha_3(n) = \lceil \log \log n \rceil$, $\alpha_4(n) = \log^* n$, $\alpha_5(n) = \lfloor \frac{1}{2} \log^* n \rfloor$, $\dots$. 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.

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 [Che86, KG92, ADD⁺93, CK93, DN94, ADM⁺95, DNS95, AS97, RS98, AWY05, CG06, DES10, SE10, Sol14, ES15, LS19]. 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 [DN94, LNS98, RS98, GLNS02, GNS05, GLNS08, HP00, MP00]. We refer the reader to the book by Narasimhan and Smid [NS07], 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) [AMS94, AM04, AWY05, CG06, DES10, KLMS21], 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 [Tho95], general metrics [Tho92] and general graphs [BGJ⁺09]. 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 [Tar79, Yao82, AS87, CR91, PD04, AWY05], and the references therein).

Alon and Schieber [AS87] and Bodlaender et al. [BTS94] 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 [AS87] 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 $\alpha_0(n) = \lceil n/2 \rceil$, $\alpha_1(n) = \lceil \sqrt{n} \rceil$, $\alpha_2(n) = \lceil \log n \rceil$, $\alpha_3(n) = \lceil \log \log n \rceil$, $\alpha_4(n) = \log^* n$, $\alpha_5(n) = \lfloor \frac{1}{2} \log^* n \rfloor$, 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. (These functions are formally defined in Section 2.1.) Bodlaender et al. [BTS94] constructed 1-spanners with diameter at most k and $O(n\alpha_k(n))$ edges, with a high running time. Solomon [Sol13] 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 [Yao82] for line metrics and later extended by Chazelle [Cha87] to general tree metrics, achieve a tradeoff of m edges versus $\Theta(\alpha(m, n))$ hop-diameter, where $\alpha(m, n)$ is the two-parameter inverse Ackermann function (see Definition 2.7). 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. [ADM⁺95] 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 Dumbbell 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 [Sol13], this gives rise to an optimal $O(n \log n)$ -time construction (in the algebraic computation tree (ACT) model¹) 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. [BFN19], which generalizes the Dumbbell Theorem of [ADM⁺95] for arbitrary doubling metrics.

1.2 Known lower bounds

The first lower bound on 1-spanners for tree metrics was given by Yao [Yao82] and it establishes a tradeoff of m edges versus hop-diameter of $\Omega(\alpha(m, n))$ for the uniform line metric. Alon and Schieber [AS87] 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 [Yao82] (see Lemma A.1), 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 [CG06], which extends that of [Yao82]: 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 [CG06] is not as basic as the uniform line metric — it is derived from hierarchically well-separated

¹Refer to Chapter 3 in [NS07] 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 [CDS01].

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 ratio² 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 [Yao82, AS87]) 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 1.1. *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 $\Omega(\frac{n}{2^{6\lceil k/2 \rceil}} \alpha_k(n))$ 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 .

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 $\log^* n$; we formally define $\alpha(n)$ in Section 2.1. 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 $2^{6\lceil k/2 \rceil}$, 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 4.1. In particular, we need to consider a more general notion of Steiner spanners³, and to prove the lower bound for a certain family of line metrics to which the uniform line metric belongs; Theorem 1.1 follows directly from Theorem 4.1. See Section 2 for the definitions.

For constant values of k , Theorem 1.1 strengthens the lower bound shown by [AS87], which applies only to stretch 1, whereas our tradeoff holds for arbitrary stretch. Whether or not the term $\frac{1}{2^{6\lceil k/2 \rceil}}$ in the bound on the number of edges in Theorem 1.1 can be removed is left open by our work. As mentioned before, we show in Appendix A that this tradeoff implies the tradeoff by [Yao82] (for stretch 1) and [CG06] (for larger stretch).

1.4 Proof overview

The starting point of our lower bound argument is the one by [AS87] for stretch 1, which applies to the uniform line metric, $U(n)$. The argument of [AS87] crucially relies on the following *separation* property: for any 4 points i, j, i', j' in $U(n)$ such that $i \leq j < i' \leq j'$, any 1-spanner path between i and j is completely disjoint to any 1-spanner path between i' and j' . Thus, for any two subintervals that do not overlap — and are thus separated in the most basic sense, any 1-spanner path between a pair of points in one of the sub-intervals does not overlap any 1-spanner path in the other sub-interval. With this separation property in mind, a rather straightforward inductive argument can be employed. Consider for concreteness the easiest case $k = 2$. The number of spanner edges satisfies the recurrence $T_2(n) = 2T_2(n/2) + \Omega(n)$. Indeed, by induction, the number of spanner edges for the left $n/2$ points is at least $T_2(n/2)$, and the same goes for the right $n/2$ points, and importantly the corresponding sets of edges are disjoint by the separation property. All is left is to reason that there are $\Omega(n)$ *cross edges*, which are the edges whose one endpoint belongs to the left $n/2$ points and the other to the right $n/2$ points; the separation property implies that the cross edges are disjoint to the aforementioned edge sets. Alas, when the stretch grows from 1 to $1 + \epsilon$, even for a tiny $\epsilon > 0$, the basic separation property no longer holds, so it is possible to have significant overlaps between the edge sets corresponding to the recursive

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

³A 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.

calls of the left half and right half of the points, and between them and the cross edges. Of course, as k increases, such issues regarding overlaps become even more challenging. To overcome these issues, [CG06] resorted to a different metric space, which is far from a uniform line metric to an extent that renders such issues of overlapping negligible. As mentioned before, the resulting metric of [CG06] has a super-polynomial aspect ratio whenever ϵ is sufficiently small or sufficiently large. Our argument, on the other hand, applies to the uniform line metric; coping with overlaps issues in this case is nontrivial. Next, we highlight some of the challenges that our argument overcomes.

The easiest case is $k = 2$. We view the line metric $U(1, n)$ as an interval $[1, n]$ of all integers from 1 to n , and also consider sub-intervals $[i, j]$ of all integers from i to j , for all $1 \leq i < j \leq n$. Our first observation is that one can achieve a similar separation property as in the case of stretch 1 by, roughly speaking, restricting the attention to a subset of $\Omega(n)$ points that are sufficiently far from the boundaries of sub-intervals. More concretely, for $\epsilon = 1/2$, consider the $n/2$ points closest to $n/2$ (i.e., the sub-interval $[n/4, 3n/4]$): While these $n/2$ points still induce $\Omega(n)$ cross edges, note that any $(1 + \epsilon)$ -spanner path between any pair of such points is contained in $[1, n]$. To achieve the required separation property, we would like to apply this observation inductively. However, a naive application is doomed to failure. Indeed, if we used the induction hypothesis only on the points inside $[n/4, 3n/4]$, i.e., on the points of $[n/4, n/2]$ as the left interval and the points of $[n/2 + 1, 3n/4]$ as the right interval, the number of cross edges would degrade by a factor of 2. Losing a factor of 2 at each recursion level cannot give a lower bound larger than $\Omega(n)$. Instead, we apply the induction hypothesis on all $n/2$ points of the left half and on all $n/2$ points of the right half, but restrict the attention to the $n/4$ points that are closest to $n/4$ in the left half and to the $n/4$ points that are closest to $3n/4$ in the right half, which gives us $n/2$ points in total, just as with the first recursion level. In this way at each level of recursion we restrict the attention to a different set of points, but of the same size $n/2$, and in this way avoid the loss. The case $k = 3$ is handled similarly, but is more intricate — primarily since the basic argument of [AS87] (for stretch 1) for $k = 3$ is more intricate than for $k = 2$. We refer the reader to the proof of Lemma 3.1 (resp. Lemma 3.3) for the proof of the case $k = 2$ (resp. $k = 3$).

For $k \geq 4$, the argument is considerably more involved. As with [AS87], we divide the interval $[1, n]$ into consecutive sub-intervals of size $\alpha_{k-2}(n)$ — the goal is to show that the number of spanner edges satisfies the recurrence $T_k(n) = (n/\alpha_{k-2}(n)) \cdot T_k(\alpha_{k-2}(n)) + \Omega(n)$, implying that $T_k(n) = \Omega(n\alpha_k(n))$. Using the separation property as in the case $k = 2$, employing the induction hypothesis on each of the $n/\alpha_{k-2}(n)$ subintervals yields the term $(n/\alpha_{k-2}(n)) \cdot T_k(\alpha_{k-2}(n))$. The crux is to show that there are $\Omega(n)$ cross edges, now defined as edges with endpoints in different sub-intervals. To ensure that the separation property holds, as in the case $k = 2$, we consider only the $n/2$ points of $[n/4, 3n/4]$. We distinguish between *global* and *non-global* points: A point is called global if it is incident on at least one cross edge, and non-global otherwise. If a constant fraction of points in $[n/4, 3n/4]$ are global, we must have $\Omega(n)$ cross edges by definition. The complementary case, where a constant fraction of points in $[n/4, 3n/4]$ are non-global, is the interesting one. We'd like to use the induction hypothesis for $k - 2$, as in [AS87], to reason that the number of cross edges induced by the non-global points is at least $T_{k-2}(n/\alpha_{k-2}(n)) = \Omega(n)$. The problem is that the non-global points induce a line metric that is *not uniform*, hence we cannot apply the induction hypothesis — overcoming this obstacle is the key difficulty in our argument. To be able to apply the induction hypothesis, we prove all the results on a generalized line metric. Consider first what we call a *t-sparse line metric*, which contains a single point in every consecutive sub-interval of size t . Such a line metric would suffice for applying the induction hypothesis (with hop-diameter $k - 2$), *if all sub-intervals inside $[n/4, 3n/4]$ contained at least one non-global point*. In this special case, we could apply the induction hypothesis on an $\alpha_{k-2}(n)$ -sparse line metric, and if we're always in this special case, we'll always have a t -sparse line metric for a growing parameter t , and it is not difficult to show that the induction step will work fine. Alas, we only know that there is a constant fraction of points around the middle of the interval and the intervals containing them form a subspace of a t -sparse line metric, which is possibly very different than this special case; thus, in general, we are unable to apply the induction hypothesis in such a way. To overcome this hurdle, we prove a stronger, more general lower bound, which concerns subspaces of t -sparse line metrics, where a constant fraction of the points is missing. On the bright side, such generalized spaces provide the required flexibility for carrying out the induction step. On the negative side, each invocation of the induction hypothesis with a smaller value of k incurs some multiplicative loss in the number of considered points, yielding the exponential on k slack in our lower bound. We refer the reader to the proof of Theorem 4.1.

2 Preliminaries

Definition 2.1 (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 2.2 (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 2.3 (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 ($[l, l+t]$, $[l+t+1, l+2t]$, ...) 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.

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

Definition 2.4 (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 in $U((l, r), t)$ 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 2.1 (Separation property). *Let $l, r, t \in \mathbb{N}$, $l \leq r$, $t \geq 1$ and let $i := \lceil \frac{1+\epsilon/2}{1+\epsilon}l + \frac{\epsilon/2}{1+\epsilon}r \rceil$, and $j := \lfloor \frac{\epsilon/2}{1+\epsilon}l + \frac{1+\epsilon/2}{1+\epsilon}r \rfloor$. 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 $[l, r]$.*

Proof. Consider a spanner path between a and b which contains a point q outside $U((l, r), t)$ such that $q > r$. (A similar argument holds for $q < l$.) The length of any such path is at least $(b - a) + 2(q - b)$. Since $b \leq j < r < q$, it holds that $2(q - b) > 2(r - j) = \frac{\epsilon}{1+\epsilon}(r - l)$. On the other hand, since $i \leq a < b \leq j$, the distance between a and b is at most $b - a \leq j - i = \frac{1}{1+\epsilon}(r - l)$. The last two inequalities imply that $2(q - b) > \epsilon(b - a)$. It follows that the spanner path between a and b containing q is of length greater than $(1 + \epsilon)(b - a)$, i.e., it has a stretch bigger than $(1 + \epsilon)$. $\square$

Corollary 2.1. *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 $[1, N]$.*

2.1 Ackermann functions

Following standard notions [Tar75, AS87, Cha87, NS07, Sol13], we will introduce two very rapidly growing functions $A(k, n)$ and $B(k, n)$, which are variants of Ackermann's function. Later, we also introduce several inverses and state their properties that will be used throughout the paper.

Definition 2.5. For all $k \geq 0$, the functions $A(k, n)$ and $B(k, n)$ are defined as follows:

$$\begin{aligned} A(0, n) &:= 2n, \text{ for all } n \geq 0, \\ A(k, n) &:= \begin{cases} 1 & \text{if } k \geq 1 \text{ and } n = 0 \\ A(k-1, A(k, n-1)) & \text{if } k \geq 1 \text{ and } n \geq 1 \end{cases} \\ B(0, n) &:= n^2, \text{ for all } n \geq 0, \\ B(k, n) &:= \begin{cases} 2 & \text{if } k \geq 1 \text{ and } n = 0 \\ B(k-1, B(k, n-1)) & \text{if } k \geq 1 \text{ and } n \geq 1 \end{cases} \end{aligned}$$

We now define the functional inverses of $A(k, n)$ and $B(k, n)$.

Definition 2.6. For all $k \geq 0$, the function $\alpha_k(n)$ is defined as follows:

$$\begin{aligned} \alpha_{2k}(n) &:= \min\{s \geq 0 : A(k, s) \geq n\}, \text{ for all } n \geq 0, \text{ and} \\ \alpha_{2k+1}(n) &:= \min\{s \geq 0 : B(k, s) \geq n\}, \text{ for all } n \geq 0, \text{ for all } n \geq 0. \end{aligned}$$

For technical convenience we define $\log x = 0$ for any $x \leq 0$, $x \in \mathbb{R}$. All the logarithms are with base 2. It is not hard to verify that $\alpha_0(n) = \lceil n/2 \rceil$, $\alpha_1(n) = \lceil \sqrt{n} \rceil$, $\alpha_2(n) = \lceil \log n \rceil$, $\alpha_3(n) = \lceil \log \log n \rceil$, $\alpha_4(n) = \log^* n$, $\alpha_5(n) = \lfloor \frac{1}{2} \log^* n \rfloor$, etc. We will use the following property of $\alpha_k(n)$.

Lemma 2.2 (cf. Lemma 12.1.16. in [NS07]). *For each $k \geq 1$, we have:*

$$\begin{aligned} \alpha_{2k}(n) &= 1 + \alpha_{2k}(\alpha_{2k-2}(n)), \text{ for all } n \geq 2, \text{ and} \\ \alpha_{2k+1}(n) &= 1 + \alpha_{2k+1}(\alpha_{2k-1}(n)), \text{ for all } n \geq 3. \end{aligned}$$

Finally, another functional inverse of $A(\cdot, \cdot)$ is defined as in [Tar75, Cha87, CG06, NS07, Sol13].

Definition 2.7 (Inverse Ackermann function). $\alpha(m, n) = \min\{i | i \geq 1, A(i, 4\lceil m/n \rceil) > \log_2 n\}$.

Finally, for all $n \geq 0$, we introduce the Ackermann function as $A(n) := A(n, n)$, and its inverse as $\alpha(n) = \min\{s \geq 0 : A(s) \geq n\}$. In [NS07], it was shown that $\alpha_{2\alpha(n)+2}(n) \leq 4$. We observe that $\alpha(n)$ satisfies $\alpha(n) \leq \log^* n$ for any $n \geq 2$.

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

In this section, we prove the lower bound for cases $k = 2$ (Lemma 3.1 in Section 3.1) and $k = 3$ (Lemma 3.3 in Section 3.2). In fact, we prove more general statements (Theorems 3.1 and 3.2), 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 4.1).

3.1 Hop diameter 2

Theorem 3.1. *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 $T_2^h(n) \geq \frac{n}{256} \cdot \alpha_2(n)$ $((l, r), t)$ -global edges which have both endpoints inside $[l, r]$.*

Remark 3.1. Recall that we consider Steiner spanners, which could possibly contain additional Steiner points from the uniform line metric.

Remark 3.2. Theorem 3.1 can be extended to $\epsilon > 1/2$. The only required change in the proof is to decrease the lengths of intervals by a factor of $1 + \epsilon$, as provided by Lemma 2.1; it is readily verified that, as a result, the lower bound decreases by a factor of $\Theta(\epsilon)$.

The theorem is proved in three steps. First, we prove Lemma 3.1, which concerns uniform line metrics. Then, we prove Lemma 3.2 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 3.1. *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 $n < 128$. 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 .

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 $T_2(n) = 2T_2(\lfloor n/2 \rfloor) + \Omega(n)$ 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 2.1, 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. Without loss of generality, assume that $|L| \leq |R|$. Suppose for contradiction that the spanner contains less than $|L|$ cross edges. This means that at least one point in $x \in L$ is not connected via a direct edge to any point on the right. Observe that, for every point $r \in R$, the 2-hop spanner path between x and r must be of the form (x, l_r, r) for some point l_r in the left set. It follows that every $r \in R$ induces a different cross edge (l_r, r) . 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:

$$T_2(n) = 2T_2(\lfloor n/2 \rfloor) + \frac{11n}{64} \stackrel{\text{induction}}{\geq} 2 \cdot \frac{\lfloor n/2 \rfloor}{16} \log \lfloor n/2 \rfloor + \frac{11n}{64} \geq \frac{n}{16} \cdot \log n ,$$

as claimed. $\square$

Lemma 3.2. *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 $[l, r]$.*

Proof. Recall the recurrence used in Lemma 3.1, $T_2(n) = 2T_2(\lfloor n/2 \rfloor) + \frac{11n}{64}$, which provides a lower bound on the number of edges of any $(1 + \epsilon)$ -spanner with hop-diameter 2 for $U(l, r)$ for any l and r such that $r = l + n - 1$. The base case for this recurrence occurs when the considered interval contains less than 128 points. Consider the recursion tree of $T_2(n)$ 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 j th interval of the i th recursion level and by $e_{i,j}$ the number of cross edges contributed by this interval. By definition, we have

$$T_2(n) = \sum_{i=0}^{\ell} \sum_{j=1}^{c_i} e_{i,j} \geq \frac{n}{16} \cdot \log n.$$

Recursion tree of $T_2(n)$ contains one node at depth $i = 0$ which corresponds to an interval of n points, two nodes at depth $i = 1$ corresponding to intervals of $\lfloor n/2 \rfloor$ points, etc. Letting n_i denote the total size

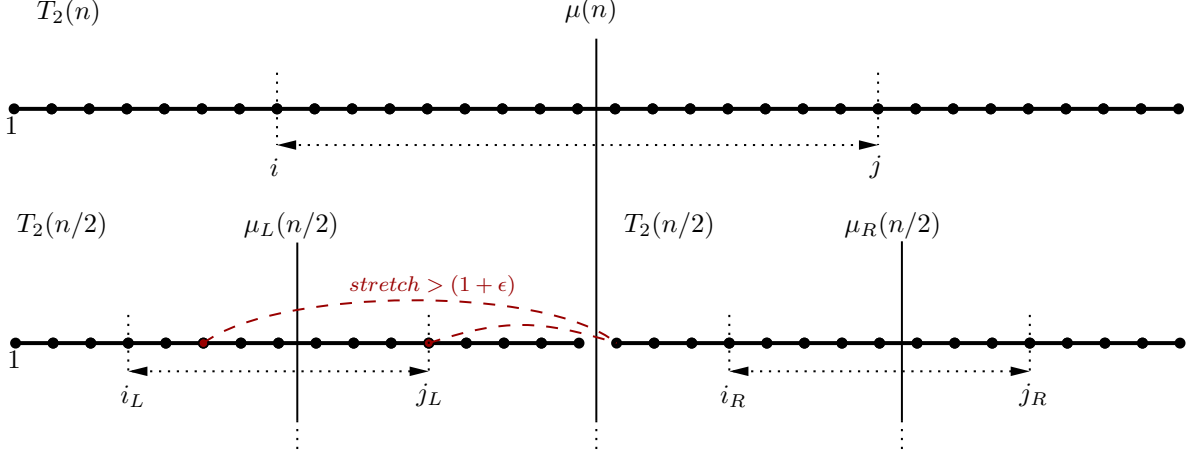


Figure 1: An illustration of the first two levels of the recurrence for the lower bound for $k = 2$ and $\epsilon = 1/2$. We split the interval $U(1, n)$ into two disjoint parts. In Lemma 3.1, we show that there will be at least $\Omega(n)$ cross edges, which are the spanner edges having endpoints in both sets. In other words, they are crossing the middle line $\mu(n)$. The values i_L and j_L are set according to Corollary 2.1 so that the spanner edges crossing $\mu(n)$ cannot be used for the left set; otherwise the resulting stretch will be bigger than $1 + \epsilon$. For example, the path formed by the dashed red edges have total length of 16 which is more than $3/2$ (i.e. more than $1 + \epsilon$) the distance between the two points they connect (which is 6). Similarly, i_R and j_R are such that the edges considered for the right set do not cross $\mu(u)$. In conclusion, the edges for the left set, right set and the cross edges will be mutually disjoint. Thus, the recurrence for the lower bound is $T_2(n) = 2T_2(\lfloor n/2 \rfloor) + \Omega(n)$, which resolves to $T_2(n) = \Omega(n \log n)$.

of intervals corresponding to nodes at depth i of recursion tree, we get

$$n_i = 2^i \cdot \left\lfloor \frac{n}{2^i} \right\rfloor \geq n - 2^i + 1.$$

The base case of the recurrence occurs whenever the considered interval contains less than 128 points. In other words, leaves of the recursion tree have $n_{i,j} < 128$. Since the size of all the intervals at depth i is the same and it equals to $\lfloor n/2^i \rfloor$, it follows that the leaves of the tree are at some level i which satisfies $64 \leq \lfloor n/2^i \rfloor < 127$. At that level, $n_i \geq 0.98 \cdot n$. For some level i of the recursion tree, we refer to all the $n - n_i$ points which are not contained in any of the intervals of that level as *ignored points*. Denote the collection of intervals containing ignored points by *ignored intervals*.

Let H' be any $(1 + \epsilon)$ -spanner on X with hop-diameter 2. 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 $U(l, r)$. Hence, at each level of recursion, instead of n points, there are at least $31n/32$ points in all the intervals of that level. We call the j th interval in the i th level *good* if it contains at least $15n_{i,j}/16$ points from X . (Recall that we have used $n_{i,j}$ to denote the number of points from $U(l, r)$ in the j th interval of the i th level.) From the definition of good interval and the fact the intervals of each recursion level contain together at least $31n/32$ points, it follows that there are at least $n/2$ points contained in the good intervals at the i th level. If more than $n/2$ points were contained in bad intervals, and any such interval had at least $1/16$ -fraction of points missing, it means that we would have more than $n/32$ points missing, which contradicts that we have at least $31n/32$ points in X . Recalling that we have at most $0.02n$ ignored points, we conclude that there is at least $0.48n$ points contained in the good intervals which were not ignored.

In Lemma 3.1, we worked on the uniform line metric $U(1, n)$ and considered two intervals L and R . We lower bounded the number of cross edges in this interval by $\min(|L|, |R|) \geq n/4 - 2 \geq 11n/64$, which holds for all $n \geq 26$. Observe that if the interval $[1, n]$ contained at least $15n/16$ points, rather than n , the same bound on the number of cross edges would hold, since $\min(|L|, |R|) - n/16 \geq n/4 - 2 - n/16 \geq 11n/64$, for all $n \geq 128$. This means that every good interval of size $n_{i,j}$ contributes at least $11n_{i,j}/64$ cross edges, as long as it is not considered as the base case (i.e. as long as $n_{i,j} \geq 128$). The same reasoning is applied inductively, so it holds also for $n_{i,j}$, rather than n , for any i, j .

Next, we analyze the recurrence $T'_2(n)$ representing the contribution of the intervals to the number of spanner edges for X . A lower bound on $T'_2(n)$ will provide a lower bound on the number of spanner edges of any spanner H' of X as defined above. Denote by Γ_i all the good intervals in the i th level which were not ignored and by $e'_{i,j}$ the number of cross edges contributed by the j th interval in the i th level. Then we have

$$T'_2(n) = \sum_{i=0}^{\ell} \sum_{j=1}^{c_i} e'_{i,j} \geq \sum_{i=0}^{\ell} \sum_{j \in \Gamma_i} e_{i,j} \geq 0.48 \cdot T_2(n) \geq 0.48 \cdot \frac{n}{16} \log n$$

as claimed. This completes the proof of Lemma 3.2. $\square$

Completing the proof of Theorem 3.1. 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 . 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 in Lemma 3.2 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. $\square$

3.2 Hop diameter 3

Theorem 3.2. *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 $[l, r]$.*

Remark 3.3. Recall that we consider Steiner spanners, which could possibly contain additional Steiner points from the uniform line metric.

Remark 3.4. Theorem 3.2 can be extended to $\epsilon > 1/2$. The only required change in the proof is to decrease the lengths of intervals by a factor of $1 + \epsilon$, as provided by Lemma 2.1; it is readily verified that, as a result, the lower bound decreases by a factor of $\Theta(\epsilon)$.

The theorem is proved in three steps. First, we prove Lemma 3.3, which concerns uniform line metrics. Then, we prove Lemma 3.4 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 3.3. *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 $n < 128$. We have that $\frac{n}{40} \log \log n < n - 1$, which is a trivial lower bound on the number of edges of H .

We now assume that $n \geq 128$. Divide 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

$$T_3(n) = \left\lfloor \frac{n}{\lfloor \sqrt{n} \rfloor} \right\rfloor \cdot T_3(\lfloor \sqrt{n} \rfloor) + \Omega(n),$$

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 j th 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 2.1, 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 3.1. Suppose that among points inside interval $[n/4, 3n/4]$, m of them are global. Then, they induce at least $m/2$ spanner edges.

Proof. Each global point contributes at least one cross edge and each edge is counted at most twice. $\square$

Claim 3.2. 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 2.1, 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. $\square$

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 3.1 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 3.2 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:

$$T_3(n) \geq \left\lfloor \frac{n}{\lfloor \sqrt{n} \rfloor} \right\rfloor \cdot T_3(\lfloor \sqrt{n} \rfloor) + \frac{n}{18} \stackrel{\text{induction}}{\geq} \lfloor \sqrt{n} \rfloor \cdot \frac{\lfloor \sqrt{n} \rfloor}{40} (\log \log \lfloor \sqrt{n} \rfloor) + \frac{n}{18} \geq \frac{n}{40} \log \log n,$$

as claimed. $\square$

Lemma 3.4. 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 $[l, r]$.

Proof. Recall the recurrence used in Lemma 3.3, $T_3(n) = \lfloor n/\lfloor \sqrt{n} \rfloor \rfloor \cdot T_3(\lfloor \sqrt{n} \rfloor) + \frac{n}{18}$, which provides a lower bound on the number of edges of any $(1 + \epsilon)$ -spanner with hop-diameter 3 for $U(l, r)$ for any l and r such that $r = l + n - 1$. The base case for this recurrence occurs whenever the considered interval contains less than 128 points. Consider the recursion tree of $T_3(n)$ 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 j th interval of the i th recursion level and by $e_{i,j}$ the number of cross edges contributed by this interval. By definition, we have

$$T_3(n) = \sum_{i=0}^{\ell} \sum_{j=1}^{c_i} e_{i,j} \geq \frac{n}{40} \cdot \log \log n$$

Recursion tree of $T_3(n)$ contains one node at depth $i = 0$ which corresponds to an interval of n points, $\lfloor n/\lfloor \sqrt{n} \rfloor \rfloor$ nodes at depth $i = 1$ corresponding to intervals of $\lfloor \sqrt{n} \rfloor$ points, etc. Letting n_i denote the total size of intervals corresponding to nodes at depth i of the recursion tree, we get

$$n_i \geq n - \sum_{j=1}^{i-1} n^{1-1/2^j}.$$

Since the size of all the intervals at depth i is the same and it equals to $\lfloor n^{1-1/2^i} \rfloor$, it follows that the leaves of the tree are at some level i which satisfies $11 \leq \lfloor n^{1-1/2^i} \rfloor < 127$. At that level, $n_i \geq 0.68 \cdot n$. For some level i of the recursion tree, we refer to all the $n - n_i$ points which are not contained in any of the intervals of that level as *ignored points*. Denote the collection of intervals containing ignored points by *ignored intervals*.

Let H' be any $(1 + \epsilon)$ -spanner on X with hop-diameter 3. 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 $U(l, r)$. Hence, at each level of recursion, instead of n points, there are at least $127n/128$ points in all the intervals of that level. We call the j th interval in the i th level *good* if it contains at least $63n_{i,j}/64$ points from X . (Recall that we have used $n_{i,j}$ to denote the number of points from $U(l, r)$ in the j th interval of the i th level.) From the definition of a good interval and the fact the intervals of each recursion level contain together at least $127n/128$ points, it follows that there are at least $n/2$ points contained in the good intervals at the i th level. If more than $n/2$ points were contained in bad intervals, and any such interval had at least $1/64$ -fraction of points missing, it means that we would have more than $n/128$ points missing; this contradicts that we have at least $127n/128$ points in X . In conclusion, at each level i , we have at least $0.18n$ points contained in the good intervals which were not ignored.

In Lemma 3.3, we worked on the uniform line metric $U(1, n)$ and considered the points inside $[n/4, 3n/4]$. We lower bounded the number of points inside $[n/4, 3n/4]$ by $15n/32$. Observe that if the interval $[1, n]$ contained at least $63n/64$ points, rather than n , the same bound on the number of cross edges would hold, for any $n \geq 128$. This means that every good interval of size $n_{i,j}$ contributes at least $n_{i,j}/18$ cross edges. The same reasoning is applied inductively, so it holds also for $n_{i,j}$ rather than n for any i, j .

Next, we analyze the recurrence $T'_3(n)$ representing the contribution of the intervals to the number of spanner edges for X . A lower bound on $T'_3(n)$ will provide a lower bound on the number of edges of any spanner H' of X as defined above. Denote by Γ_i all the good intervals in the i th level which were not ignored and by $e'_{i,j}$ the number of cross edges contributed by the j th interval in the i th level. Then, we have

$$T'_3(n) = \sum_{i=0}^{\ell} \sum_{j=1}^{c_i} e'_{i,j} \geq \sum_{i=0}^{\ell} \sum_{j \in \Gamma_i} e_{i,j} = 0.18 \cdot T_3(n) \geq 0.18 \cdot \frac{n}{40} \log \log n$$

as claimed. This completes the proof of Lemma 3.4. $\square$

Completing the proof of Theorem 3.2. 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 . 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 3.4 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. $\square$

4 Lower bound for constant hop-diameter

We proceed to prove our main result, which is a generalization of Theorem 1.1. In particular, invoking Theorem 4.1 stated below where X is the uniform line metric $U(1, n)$ gives Theorem 1.1.

Theorem 4.1. *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 $[l, r]$.*

Remark 4.1. Recall that we consider Steiner spanners, which could possibly contain additional Steiner points from the uniform line metric.

Remark 4.2. Theorem 4.1 can be extended to $\epsilon > 1/2$. The only required change in the proof is to decrease the lengths of intervals by a factor of $1 + \epsilon$, as provided by Lemma 2.1; it is readily verified that, as a result, the lower bound decreases by a factor of $\Theta(\epsilon)$.

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 3.1 and 3.2, 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 .

For the base case, we take $1 \leq n \leq 10000$. We have $\frac{n}{2^{6\lfloor k/2 \rfloor + 2}} \alpha_k(n)$, which is at most $\frac{n}{2^{6\lfloor k/2 \rfloor + 2}} \log^*(n) \leq n - 1$, a trivial lower bound on the number of the edges of any spanner.

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

$$T_k(n) = \left\lfloor \frac{n}{\alpha_{k-2}(n)} \right\rfloor \cdot T_k(\alpha_{k-2}(n)) + \Omega\left(\frac{n}{2^{3k}}\right),$$

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

There are $\lfloor n/b \rfloor$ $((1, N), bt)$ -intervals containing exactly b points. For any j such that $1 \leq j \leq \lfloor n/b \rfloor$, the j th $((1, N), bt)$ -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 $p \geq n/2 - 2\alpha_{k-2}(n)$, but we will use a weaker bound:

$$p \geq n/4. \quad (1)$$

Definition 4.1. 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

$$m \geq p'/2. \quad (2)$$

Next, we prove that

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

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 2.3 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

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

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 2.1 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 $(a, a', \dots, b', b)$, 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 :

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

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 $\lceil p''/b \rceil \geq (1 - 1/2^{k+2}) \cdot \lceil p/b \rceil$), 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

$$\frac{p - p'}{b} - 1 < \left\lfloor \frac{p - p'}{b} \right\rfloor = \frac{p''}{b} < \left(1 - \frac{1}{2^{k+2}}\right) \cdot \left\lceil \frac{p}{b} \right\rceil < \left(1 - \frac{1}{2^{k+2}}\right) \cdot \frac{p}{b} + 1.$$

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

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

For the second inequality we have used Lemma 2.2, 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.

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, $T_k(n) = \lfloor n/\alpha_{k-2}(n) \rfloor \cdot T_k(\alpha_{k-2}(n)) + \Omega(n/2^{3k})$, 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 j th

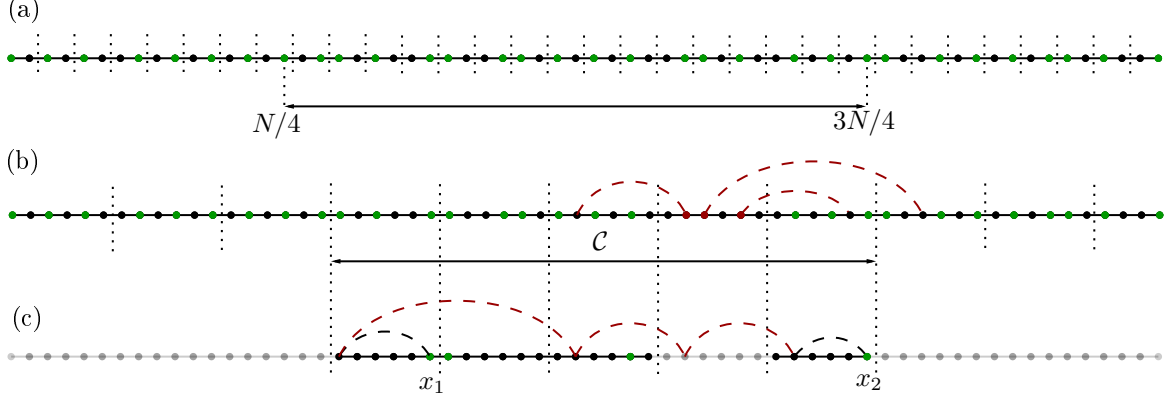


Figure 2: Constructing a new line metric and invoking the induction hypothesis. **(a)** We have $n = 32$, $k = 5$, and a 2-sparse line metric $U((1, 64), 2)$ with representatives of each $((1, 64), 2)$ -interval highlighted in green. **(b)** Since $b = \alpha_{k-2}(n) = 3$, we consider a collection of $((1, 64), 6)$ -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.

interval of the i th level and by $e_{i,j}$ the number of $((1, N), bt)$ -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 j th interval in the i th level is denoted by $e'_{i,j}$. We call the j th interval in the i th 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 j th interval of the i th 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 i th level. Denote the collection of all the good intervals at the i th level by Γ_i .

Recall that we are working with recurrence $T_k(n) = \lfloor n/\alpha_{k-2}(n) \rfloor \cdot T_k(\alpha_{k-2}(n)) + \Omega(n/2^{3k})$. 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 i th 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 $\alpha_{k-2}^{(j)}(n)$ value of $\alpha_{k-2}(\cdot)$ iterated on n , i.e. $\alpha_{k-2}^{(0)}(n) = n$, $\alpha_{k-2}^{(1)}(n) = \alpha_{k-2}(n)$, $\alpha_{k-2}^{(2)}(n) = \alpha_{k-2}(\alpha_{k-2}(n))$, etc. In general, for $i \geq 2$, we have

$$\begin{aligned} n_i &\geq n - \sum_{j=2}^i \frac{n\alpha_{k-2}^{(j-1)}(n)}{\alpha_{k-2}^{(j-2)}(n)} \\ &\geq n - n \cdot \sum_{j=2}^i \frac{\lceil \log^{(j-1)}(n) \rceil}{\lceil \log^{(j-2)}(n) \rceil}. \end{aligned}$$

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 i th level whose contribution is not ignored. Then we have

$$T'_k(n) = \sum_{i=1}^{\ell} \sum_{j=1}^{c'_i} e'_{i,j} \geq \sum_{i=1}^{\ell} \sum_{j \in \Gamma_i} e_{i,j} \geq 0.49 \cdot T_k(n) \geq \frac{n}{2^{6\lfloor k/2 \rfloor + 4}} \alpha_k(n),$$

as claimed. This concludes the proof of assertion 2.

We have thus completed the inductive step for k . □

References

- [ADD⁺93] I. Althöfer, G. Das, D. Dobkin, D. Joseph, and J. Soares. On sparse spanners of weighted graphs. *Discrete Computational Geometry*, 9(1):81–100, 1993.
- [ADM⁺95] 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.
- [AM04] I. Abraham and D. Malkhi. Compact routing on Euclidean metrics. In *Proc. of 23rd PODC*, pages 141–149, 2004.
- [AMS94] S. Arya, D. M. Mount, and M. H. M. Smid. Randomized and deterministic algorithms for geometric spanners of small diameter. In *Proc. of 35th FOCS*, pages 703–712, 1994.
- [AS87] Noga Alon and Baruch Schieber. Optimal preprocessing for answering on-line product queries. 1987.
- [AS97] S. Arya and M. H. M. Smid. Efficient construction of a bounded degree spanner with low weight. *Algorithmica*, 17(1):33–54, 1997.
- [AWY05] P. K. Agarwal, Y. Wang, and P. Yin. Lower bound for sparse Euclidean spanners. In *Proc. of 16th SODA*, pages 670–671, 2005.
- [BFN19] 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.
- [BGJ⁺09] A. Bhattacharyya, E. Grigorescu, K. Jung, S. Raskhodnikova, and D. P. Woodruff. Transitive-closure spanners. In *Proc. of 20th SODA*, pages 932–941, 2009.
- [BTS94] H. L. Bodlaender, G. Tel, and N. Santoro. Trade-offs in non-reversing diameter. *Nord. J. Comput.*, 1(1):111–134, 1994.
- [CDS01] D. Z. Chen, G. Das, and M. H. M. Smid. Lower bounds for computing geometric spanners and approximate shortest paths. *Discrete Applied Mathematics*, 110(2-3):151–167, 2001.
- [CG06] H. T.-H. Chan and A. Gupta. Small hop-diameter sparse spanners for doubling metrics. In *Proc. of 17th SODA*, pages 70–78, 2006.
- [Cha87] Bernard Chazelle. Computing on a free tree via complexity-preserving mappings. *Algorithmica*, 2:337–361, 1987.
- [Che86] L. P. Chew. There is a planar graph almost as good as the complete graph. In *Proceedings of the Second Annual Symposium on Computational Geometry*, SCG ‘86, pages 169–177, 1986.
- [CK93] P. B. Callahan and S. R. Kosaraju. Faster algorithms for some geometric graph problems in higher dimensions. In *Proc. of 4th SODA*, pages 291–300, 1993.

- [CR91] B. Chazelle and B. Rosenberg. The complexity of computing partial sums off-line. *Int. J. Comput. Geom. Appl.*, 1:33–45, 1991.
- [DES10] Y. Dinitz, M. Elkin, and S. Solomon. Low-light trees, and tight lower bounds for Euclidean spanners. *Discrete & Computational Geometry*, 43(4):736–783, 2010.
- [DN94] G. Das and G. Narasimhan. A fast algorithm for constructing sparse Euclidean spanners. In *Proc. of 10th SOCG*, pages 132–139, 1994.
- [DNS95] G. Das, G. Narasimhan, and J. Salowe. A new way to weigh malnourished euclidean graphs. In *Proceedings of the 6th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA ’95, pages 215–222, 1995.
- [ES15] Michael Elkin and Shay Solomon. Optimal euclidean spanners: Really short, thin, and lanky. *J. ACM*, 62(5):35:1–35:45, 2015.
- [GLNS02] J. Gudmundsson, C. Levcopoulos, G. Narasimhan, and M. H. M. Smid. Approximate distance oracles for geometric graphs. In *Proc. of 13th SODA*, pages 828–837, 2002.
- [GLNS08] J. Gudmundsson, C. Levcopoulos, G. Narasimhan, and M. H. M. Smid. Approximate distance oracles for geometric spanners. *ACM Transactions on Algorithms*, 4(1), 2008.
- [GNS05] J. Gudmundsson, G. Narasimhan, and M. H. M. Smid. Fast pruning of geometric spanners. In *Proc. of 22nd STACS*, pages 508–520, 2005.
- [HP00] Y. Hassin and D. Peleg. Sparse communication networks and efficient routing in the plane. In *Proc. of 19th PODC*, pages 41–50, 2000.
- [KG92] J. M. Keil and C. A. Gutwin. Classes of graphs which approximate the complete Euclidean graph. *Discrete and Computational Geometry*, 7(1):13–28, 1992.
- [KLMS21] 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.
- [LNS98] C. Levcopoulos, G. Narasimhan, and M. H. M. Smid. Efficient algorithms for constructing fault-tolerant geometric spanners. In *Proc. of 30th STOC*, pages 186–195, 1998.
- [LS19] Hung Le and Shay Solomon. Truly optimal euclidean spanners. In *FOCS*, pages 1078–1100. IEEE Computer Society, 2019.
- [MP00] Y. Mansour and D. Peleg. An approximation algorithm for min-cost network design. *DIMACS Series in Discr. Math and TCS*, 53:97–106, 2000.
- [NS07] Giri Narasimhan and Michiel H. M. Smid. *Geometric spanner networks*. Cambridge University Press, 2007.
- [PD04] M. Pătraşcu and E. D. Demaine. Tight bounds for the partial-sums problem. In *Proc. of 15th SODA*, pages 20–29, 2004.
- [RS98] S. B. Rao and W. D. Smith. Approximating geometrical graphs via “spanners” and “banyans”. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, STOC ’98, pages 540–550, 1998. Full version at <http://graphics.stanford.edu/courses/cs468-06-winter/Papers/rs-tsp.pdf>.
- [SE10] S. Solomon and M. Elkin. Balancing degree, diameter and weight in Euclidean spanners. In *Proc. of 18th ESA, Part 1*, pages 48–59, 2010.
- [Sol13] Shay Solomon. Sparse euclidean spanners with tiny diameter. *ACM Trans. Algorithms*, 9(3):28:1–28:33, 2013.
- [Sol14] Shay Solomon. From hierarchical partitions to hierarchical covers: optimal fault-tolerant spanners for doubling metrics. In *STOC*, pages 363–372. ACM, 2014.

- [Tar75] Robert E. Tarjan. Efficiency of a good but not linear set union algorithm. *Journal of the ACM*, 22(2):215–225, 1975.
- [Tar79] R. E. Tarjan. Applications of path compression on balanced trees. *J. ACM*, 26(4):690–715, 1979.
- [Tho92] M. Thorup. On shortcutting digraphs. In *Proc. of 18th WG*, pages 205–211, 1992.
- [Tho95] M. Thorup. Shortcutting planar digraphs. *Combinatorics, Probability & Computing*, 4:287–315, 1995.
- [Yao82] A. C. Yao. On constructing minimum spanning trees in k -dimensional spaces and related problems. *SIAM Journal on Computing*, 11(4):721–736, 1982.

A Tradeoff using two-parameter Ackermann function

In this appendix, we prove that the tradeoff of [AS87] of k vs. $\Omega(n\alpha_k(n))$ between hop-diameter and number of edges implies the tradeoff of [Yao82] and [CG06] of $\Omega(\alpha(m, n))$ vs. m . We also prove a variant of this result, relevant to our lower bound (Theorem 1.1) that has an exponential on k slack in the number of edges. Specifically, we prove Lemma A.1 and Corollary A.1.

We start with making the following simple claim.

Claim A.1. For every $i \geq 0$ and every $j \geq 4$, we have $A(i + 1, j) \geq A(i, 2j)$.

Proof. Observe that $A(1, j) = 2^j$. We have $A(i + 1, j) = A(i, A(i + 1, j - 1)) \geq A(i, 2^{j-1}) \geq A(i, 2j)$, whenever $j \geq 4$. $\square$

Lemma A.1. Let $\mathcal{G}$ be a graph family and let k', n_0 be a constant such that for every integer $n' \geq n_0$ any k' -hop spanner on an n' -vertex graph $G_{n'}$ from $\mathcal{G}$ has at least $n'\alpha_{k'}(n')$ edges. Then, for any choice of integers $n \geq n_0$ and $m \geq n$, any m -edge spanner on an n -vertex graph G_n from $\mathcal{G}$ has hop diameter of at least $\alpha(m, n)$.

Proof. Let H be an arbitrary m -edge spanner on G_n from $\mathcal{G}$ as in the lemma statement. If $n\alpha_0(n) \leq m$, then $\alpha(m, n)$ is a small constant, and the hop-diameter of H is trivially $\Omega(1)$. We henceforth assume that $n\alpha_0(n) > m$, hence there is a unique k such that

$$n \cdot \alpha_k(n) \leq m < n \cdot \alpha_{k-1}(n). \quad (5)$$

From the lemma statement, the hop-diameter of H is greater than $k - 1$, i.e., it is at least:

$$\begin{aligned}
k &\geq \min \{i \mid A(i, \min \{s \mid A(k, s) \geq n\}) \geq n\} \\
&\geq \min \{i \mid A(i, \alpha_{2k}(n)) \geq n\} && \text{(by definition of } \alpha_{2k}(n)\text{)} \\
&\geq \min \{i \mid A(i, \alpha_{2k}(n)) > \log n\} \\
&\geq \min \{i \mid A(i, \alpha_k(n)) > \log n\} \\
&\geq \min \left\{ i \mid A\left(i, \frac{m}{n}\right) > \log n \right\} && \text{(Equation (5))} \\
&\geq \min \left\{ i \mid A\left(i, 4 \left\lceil \frac{m}{n} \right\rceil\right) > \log n \right\} \\
&= \alpha(m, n)
\end{aligned}$$

This completes the proof of Lemma A.1. $\square$

We proceed to proving Corollary A.1, which is a variant of Lemma A.1, relevant to our lower bound (Theorem 1.1) that has an exponential on k slack in the number of edges. By Theorem 1.1, any $(1 + \epsilon)$ -spanner with hop-diameter k on any n -point uniform line metric must have at least $\frac{n}{2^{6\lceil k/2 \rceil + 4}} \alpha_k(n)$ edges. (This refined bound on the number of edges, which does not use the Ω -notation, is due to Theorem 4.1, which is a generalization of Theorem 1.1.)

Corollary A.1. *For any two positive integers m and n such that $m < n^2/32$, let k be the unique integer such that $\frac{n}{2^{6\lfloor k/2 \rfloor + 4}} \cdot \alpha_k(n) \leq m < \frac{n}{2^{6\lfloor (k-1)/2 \rfloor + 4}} \cdot \alpha_{k-1}(n)$. For any choice of $\epsilon \in [0, 1/2]$, any m -edge $(1 + \epsilon)$ -spanner for the uniform line metric with n points must have hop diameter of at least $\Omega(\alpha(m, n) - k)$.*

Proof. The value k as in the statement always exists since for $k = 0$, we have $\frac{n}{2^{6\lfloor k/2 \rfloor + 4}} \alpha_k(n) \geq \frac{n^2}{32} > m$ and the sequence $\{\frac{n}{2^{6\lfloor k/2 \rfloor + 4}} \alpha_k(n)\}_{k=0}^\infty$ is decreasing.

Similarly to the proof of Lemma A.1, we can lower bound k as follows.

$$\begin{aligned}
k &\geq \min \{i \mid A(i, \min \{s \mid A(k, s) \geq n\}) \geq n\} \\
&\geq \min \{i \mid A(i, \alpha_{2k}(n)) \geq n\} && \text{(by definition of } \alpha_{2k}(n)\text{)} \\
&\geq \min \{i \mid A(i, \alpha_{2k}(n)) > \log n\} \\
&\geq \min \{i \mid A(i, \alpha_k(n)) > \log n\} \\
&\geq \min \left\{ i \mid A\left(i, 2^{6\lfloor k/2 \rfloor + 4} \cdot \frac{m}{n}\right) > \log n \right\} \\
&\geq \min \left\{ i \mid A\left(i + 6\lfloor k/2 \rfloor + 4, \frac{m}{n}\right) > \log n \right\} && \text{(Claim A.1)} \\
&\geq \min \left\{ i \mid A\left(i, \frac{m}{n}\right) > \log n \right\} - 6\lfloor k/2 \rfloor - 4 \\
&\geq \min \left\{ i \mid A\left(i, 4\left\lceil \frac{m}{n} \right\rceil\right) > \log n \right\} - 6\lfloor k/2 \rfloor - 4 \\
&\geq \alpha(m, n) - 6\lfloor k/2 \rfloor - 4 \\
&= \Omega(\alpha(m, n) - k)
\end{aligned}$$

This completes the proof of Corollary A.1. □