Sparse Source-wise and Pair-wise Distance Preservers

Extended Abstract

Don Coppersmith *　　Michael Elkin †‡

Abstract

We introduce and study the notions of *pair-wise* and *source-wise preservers.*

Given an undirected N-vertex graph $G = (V, E)$ and a subset $\mathcal{P}$ of pairs of vertices, let $G' = (V, H)$, $H \subseteq E$, be called a *pair-wise preserver of G with respect to $\mathcal{P}$* if for every pair $\{u, w\} \in \mathcal{P}$, $dist_{G'}(u, w) = dist_G(u, w)$. For a set $\mathcal{S} \subseteq V$ of *sources*, a pair-wise preserver of G with respect to the set of all pairs $\mathcal{P} = \binom{\mathcal{S}}{2}$ of sources is called a *source-wise preserver of G with respect to $\mathcal{S}$.*

We prove that for every undirected possibly weighted N-vertex graph G and every subset $\mathcal{P}$ of $P = O(N^{1/2})$ pairs of vertices of G, there exists a *linear-size* pair-wise preserver of G with respect to $\mathcal{P}$. Consequently, for every subset $\mathcal{S} \subseteq V$ of $S = O(N^{1/4})$ sources, there exists a *linear-size* source-wise preserver of G with respect to $\mathcal{S}$. On the negative side we show that neither of the two exponents (1/2 and 1/4) can be improved even when the attention is restricted to unweighted graphs.

Our lower bounds involve constructions of dense convexly independent sets of vectors with small Euclidean norms. We believe that the link between the areas of *Discrete Geometry* and *spanners* that we establish is of independent interest, and might be useful in the study of other problems in the area of low-distortion embeddings.

1 Introduction

For a graph $G = (V, E)$, its sparse subgraph $G' = (V, H)$, $H \subseteq E$, is called a *spanner* of G if the metric space that is defined by G' is close in some respect to the metric space that is defined by G.

Graph spanners were introduced in a pioneering paper of Peleg and Schaffer [24], and since then they were used as an underlying combinatorial structure for many applications, mostly in the areas of Graph Algorithms and Distributed Computing. Among the most prominent applications of spanners are algorithms for computing almost shortest paths [2, 12, 14], routing algorithms [25, 4, 26], algorithms for constructing synchronizers [3, 5], and for network design [21]. There are also indirect applications for distance labeling and distance oracles [23, 16, 27]. Significant research effort was also invested in the problem of devising *efficient* algorithms for *constructing* spanners [2, 12, 13, 14, 9]. Despite this extensive study of the *algorithmic aspects* of spanners, so far relatively not much attention was devoted to their *combinatorial properties.* The study of these properties is the subject of the current paper.

The first result of this kind was due to Peleg and Schaffer, who have shown [24] that for any unweighted undirected N-vertex graph $G = (V, E)$, and a positive integer parameter $\kappa = 1, 2, \ldots$, there exists a subgraph $G' = (V, H)$, $H \subseteq E$, with $N^{1+O(1/\kappa)}$ edges that satisfies that for every pair of vertices $u, w \in V$, the distance between them in G', denoted $dist_{G'}(u, w)$, is at most κ times greater than the distance between them in G, $dist_G(u, w)$. It was also shown in [24] that this tradeoff is optimal up to the constants hidden by the O-notation. Next result of this flavor was due to Dor et al. [13], who have shown that for every unweighted undirected N-vertex graph $G = (V, E)$ there exists a subgraph $G' = (V, H)$, $H \subseteq E$, with $O(N^{3/2} \cdot \log N)$ edges such that for every pair of vertices $u, w \in V$, $dist_{G'}(u, w) \leq dist_G(u, w) + 2$ (such a subgraph is called an *additive* 2-*spanner*). A lower bound of $\Omega(N^{3/2})$ on the size of an additive 2-spanner follows directly from [24]. The gap of $\log N$ was closed (by improving the upper bound) in [15]. Very recently Baswana et al. [8] have proved the existence of additive 6-spanners with $O(n^{4/3})$ edges for every (undirected unweighted) graph.

Further, Elkin and Peleg [15] have shown that for every $\epsilon > 0$, $\kappa = 1, 2, \ldots$, there exists $\beta = \beta(\epsilon, \kappa)$ such that for every unweighted undirected N-vertex graph $G = (V, E)$ there exists a subgraph $G' = (V, H)$ with $O(N^{1+1/\kappa})$ edges, such that for every pair of vertices $u, w \in V$, $dist_{G'}(u, w) \leq (1 + \epsilon) \cdot dist_G(u, w) + \beta$.

Finally, recently Bollobas et al. [10] have shown that for every unweighted undirected N-vertex graph $G = (V, E)$, and a positive integer parameter D, there exists a subgraph $G' = (V, H)$ with $|H| = O(N^2/D)$ edges that preserves all the distances between pairs of vertices that are at distance D or more one from another

*IBM Research, Yorktown Heights, NY, USA, 10598; dcopper@us.ibm.com

†Department of Computer Science, Yale University, New Haven, CT, USA, 06520-8285, elkin@cs.yale.edu.

‡This work was supported by the DoD University Research Initiative (URI) administered by the Office of Naval Research under Grant N00014-01-1-0795.

in the graph G. It is also shown in [10] that this result is optimal up to a small constant factor.

In this paper we continue this line of research and introduce the following notions. Given an N-vertex graph $G = (V,E)$ and a subset $\mathcal{P}$ of pairs of vertices, let $G' = (V,H)$, $H \subseteq E$, be called a *pair-wise preserver of G with respect to $\mathcal{P}$* if for every pair $\{u,w\} \in \mathcal{P}$, $dist_{G'}(u,w) = dist_G(u,w)$. For a set $\mathcal{S} \subseteq V$ of vertices, called *sources*, a pair-wise preserver of G with respect to the set of all pairs $\mathcal{P} = \binom{\mathcal{S}}{2}$ of sources is called a *source-wise preserver of G with respect to $\mathcal{S}$*. We prove the following results.

THEOREM 1.1. *1. For every undirected weighted or unweighted N-vertex graph $G = (V,E)$, and a subset $\mathcal{P}$ of $P = O(N^{1/2})$ pairs of vertices of G, there exists a* linear-size *pair-wise preserver $G' = (V,H)$, $H \subseteq E$, of G with respect to the set $\mathcal{P}$. Consequently, for every graph G and subset $\mathcal{S} \subseteq V$ of $S = O(N^{1/4})$ vertices, there exists a* linear-size *source-wise preserver $G' = (V,H)$, $H \subseteq E$, of G with respect to the set $\mathcal{S}$.*

2. *For every value α, $1/2 < \alpha < 2$, there exist infinitely many values of N and $P = \Theta(N^\alpha)$ for which there exist unweighted undirected N-vertex graphs $G = (V,E)$, and subsets $\mathcal{P}$ of P pairs of vertices of G such that any pair-wise preserver $G' = (V,H)$, $H \subseteq E$, of G with respect to the set $\mathcal{P}$ contains $\omega(N+P)$ edges.*

3. *For every value α, $1/4 < \alpha < 9/16$, there exist infinitely many values of N and $S = \Theta(N^\alpha)$ for which there exist unweighted undirected N-vertex graphs $G = (V,E)$, and subsets $\mathcal{S} \subseteq V$ of S sources such that any source-wise preserver $G' = (V,H)$, $H \subseteq E$, of G with respect to the set $\mathcal{S}$ contains $\omega(N+S^2)$ edges.*

Note that the lower bounds (assertions 2 and 3) apply even for *unweighted undirected* graphs.

We remark that these results are special cases of far more general theorems that we prove. The detailed exposition of the latter theorems is deferred to Section 1.2.

1.1 Motivation We believe that the problem of constructing sparse pair-wise and source-wise preservers is an important basic combinatorial problem. Our results on it enable to get a deep insight into the metric properties of graphs. Further, we believe that the ultimate resolution of this problem is essential for resolving other major open problems in the area of spanners, such as the question of existence or nonexistence of sparse additive spanners (see [13, 10]).

From a broader perspective, spanners are currently widely recognized as one of the topics of the area of Low-Distortion Embeddings (see, e.g., the section on spanners in the recent survey paper by Indyk and Matousek [18]). The latter area is currently one of the most intensively studied subdisciplines of Theoretical Computer Science. For many fundamental existential results in this area (such as Bourgain's embeddings [11], Johnson-Lindenstrauss dimension reduction [20], average-stretch tree embeddings of Alon et al. [1]) were found multiple important algorithmic applications, sometimes many years after these existential results were proven. We believe that our research of pair-wise and source-wise preservers will also bear algorithmic fruits.

We remark that in this paper we do not explore the *algorithmic* potential of the pair-wise and source-wise preservers, but focus instead on their *combinatorial* properties. We hope that this potential will be fully explored in the consequent work. Another promising direction that we did not study is the *approximate* variants of the pair-wise and source-wise preservers. In our opinion it is very likely that these approximate variants will be useful for the design of improved algorithms for fast distance estimation. However, we feel that the study of these approximate variants would have been premature before gaining thorough understanding of the *exact* variants of these problems. We believe that this paper is a major step towards achieving such an understanding.

1.2 Our Results The specific behavior of our lower bound on the size of pair-wise preservers for unweighted undirected graphs is parameterized by the "dimension" parameter d, and has the following form: for $d = 2,3,\ldots$, for $\Omega\left(N^{2\cdot\frac{d^2-d-1}{(d-1)(d+2)}}\right) = P = O\left(N^{2\cdot\frac{d^2+d-1}{d(d+3)}}\right)$, the lower bound is $|H| = \Omega\left(N^{\frac{2d}{d^2+1}} \cdot P^{\frac{d(d-1)}{d^2+1}}\right)$. (Observe that if we denote $f(d) = 2\cdot\frac{d^2-d-1}{(d-1)(d+2)}$ then the condition on P is of the form $\Omega(N^{f(d)}) = P = O(N^{f(d+1)})$.) Note that this result implies directly assertion 2 of Thm. 1.1, that is, this lower bound is *super-linear in $N+P$* for the entire feasible range of P, $\omega(\sqrt{N}) = P = o(N^2)$. For undirected *weighted* graphs we show an even stronger lower bound of $|H| = \Omega((N \cdot P)^{2/3})$, and this lower bound is also *super-linear in $N+P$* in the same range of P.

We also show that there are unweighted undirected N-vertex graphs $G = (V,E)$, and subsets $\mathcal{S} \subseteq V$ of S vertices, such that any *source-wise* preserver $G' = (V,H)$ of G with respect to $\mathcal{S}$ contains $|H| = \Omega(\max\{N^{9/11} \cdot S^{6/11}, N^{10/11} \cdot S^{4/11}\})$ edges. This lower bound is *super-linear in* $N+S^2$ for $\omega(N^{1/4}) = S = o(N^{9/16})$ (i.e., it implies assertion 3 of Thm. 1.1). This result cannot be extended for $S = O(N^{1/4})$ in view of our upper bound of Thm. 1.1(1). For undirected *weighted* graphs we show a slightly stronger lower bound of $|H| = \Omega(N^{6/7} \cdot S^{4/7})$. This lower bound is *super-linear in* $N+S^2$ in a slightly wider range $\omega(N^{1/4}) = S = o(N^{3/5})$.

Finally, we show two upper bounds. First, we show that for every undirected possibly weighted N-vertex graph $G = (V, E)$, and a subset $\mathcal{P}$ of P pairs of vertices, there exists a pair-wise preserver $G' = (V, H)$ of G with respect to $\mathcal{P}$ with $|H| = O(N + \sqrt{N} \cdot P)$ edges. Note that this upper bound implies assertion 1 of Thm. 1.1. Second, we show an analogous upper bound of $|H| = O(\sqrt{P} \cdot N)$ that applies even to the most general case of weighted directed graphs. See Tables 1 and 2 for the summary of these results.

Note that our results do not rule out the possibility that for any unweighted undirected N-vertex graph $G = (V, E)$ and any subset $\mathcal{S}$ of $S = \Omega(N^{9/16})$ vertices, there exists a source-wise preserver $G' = (V, H)$, $H \subseteq E$, of G with respect to $\mathcal{S}$ with only $O(S^2)$ edges. To prove or disprove this statement is a challenging open problem. Also, there are some gaps between the curves of the upper and lower bounds. Closing these gaps is a very interesting open problem as well. We hope that our paper will trigger future research on these fundamental problems.

1.3 Our Techniques

The lower bounds constitute the technically more involved part of the paper, and the techniques that are used for proving them are mainly from the area of Discrete Geometry.

Specifically, we consider certain subsets of points that belong to high-dimensional integer lattices, and build graphs whose sets of vertices are those subsets. Next, for each of those points we build a convex polytope whose extreme points are also vertices of the graph, and connect this point to all the extreme points of its polytope via edges. This way the edgesets of our graphs are constructed. In some cases additional vertices and edges that have no geometric interpretation are added to the graph, and the proofs combine geometric and combinatorial techniques.

The main polytope that is used for constructing the edgesets is the convex hull of the set of integer points of a ball with a large radius $R \gg 1$. This polytope is an important object of study in Discrete Geometry, and our proofs make use of the most recent advances in the study of this object. Specifically, we use the results of Barany and Larman [7] and Balog and Barany [6] that analyze the number of vertices and faces of this polytope. We believe that introducing the techniques from the area of Discrete Geometry to the study of spanners is our important technical contribution.

To motivate the use of geometric (Euclidean) graphs, we remark that we are not aware of other constructions of *dense* graphs in which the structure of shortest paths is well-understood and relatively simple. Designing significantly simpler constructions of graphs with these properties is a challenging open problem.

The proofs of our upper bounds are conceptually simpler, and are based on double-counting of appropriate combinatorial quantities.

Structure of the paper: Section 2 is devoted to the lower bound on the cardinality of pair-wise preservers for weighted graphs. In Section 3 we turn to the lower bounds for unweighted graphs. In Section 4 we describe our lower bound for source-wise preservers for unweighted graphs. Due to space limitations, the section that contains the upper bounds, as well as most proofs from the sections that are devoted to the lower bounds, is omitted from this extended abstract.

Preliminaries

A sequence $\Pi = (v_0, v_1, \ldots, v_\ell)$ of distinct vertices of the graph $G = (V, E)$ is called a *path* if $\{v_i, v_{i+1}\} \in E$ (or $(v_i, v_{i+1}) \in E$ if G is a directed graph) for every index $i \in [(\ell)] = \{0, 1, \ldots, \ell - 1\}$. The *length* of the path Π, denoted $L(\Pi)$, is defined as ℓ if the graph is unweighted, as $\sum_{i=1}^{\ell} \omega(\{v_{i-1}, v_i\})$ if it is weighted. For a pair of vertices $u, w \in V$, let $\hat{\Pi}_{u,w}$ denote the set of paths between u and w in G (or from u to w if G is a directed graph).

The *convex hull* of the set $\{\gamma_1, \ldots, \gamma_k\} \subset \mathbb{R}^d$ of vectors is the set of all convex combinations of these vectors. A *polytope* is the convex hull of some set of vectors. It is known that a polytope can also be defined as the set of feasible solutions of n linear inequalities, $n > d$, and these two definitions are equivalent (see, e.g., [22]). A *polygon* is a two-dimensional polytope.

The set $\{\gamma_1, \ldots, \gamma_k\} \subset \mathbb{R}^d$ is called a *convexly independent set* (henceforth, *CIS*), if for every index $i \in [k] = \{1, 2, \ldots, k\}$, the vector γ_i cannot be expressed as a convex combination of the vectors $\gamma_1, \ldots, \gamma_{i-1}, \gamma_{i+1}, \ldots, \gamma_k$.

The *extreme points* or *vertices* are 0-faces of the polytope, and the $(d-1)$-faces are called the *facets* of the polytope. Equivalently, an extreme point of the polytope P can be defined as a vector $v \in P$ that cannot be expressed as a convex combination of other vectors of P.

	Undirected unweighted graphs	Undirected weighted graphs	Directed weighted graphs
U.b.	$O(\min\{N \cdot \sqrt{P}, \sqrt{N} \cdot P\})$	$O(\min\{N \cdot \sqrt{P}, \sqrt{N} \cdot P\})$	$O(N \cdot \sqrt{P})$
L.b.	$\Omega\left(\max\{N^{\frac{2d}{d^2+1}} \cdot P^{\frac{d(d-1)}{d^2+1}} : d = 2, 3, \ldots\}\right)$	$\Omega((NP)^{2/3})$	$\Omega((NP)^{2/3})$

Table 1: A summary of our results on pair-wise preservers. The columns correspond to different types of graphs. The upper (resp., lower) bounds appear in the first (resp., second) row.

	Undirected unweighted graphs	Undirected weighted graphs
U.b.	$O(\min\{\sqrt{N} \cdot S^2, N \cdot S\})$	$O(\min\{\sqrt{N} \cdot S^2, N \cdot S\})$
L.b.	$\Omega\left(\max\{N^{9/11} \cdot S^{6/11}, N^{10/11} \cdot S^{4/11}\}\right)$	$\Omega(N^{6/7} \cdot S^{4/7})$

Table 2: A summary of our results on source-wise preservers.

2 Lower Bounds: Pair-wise Preservers for Weighted Graphs

In this section we show lower bounds on the cardinalities of pair-wise preservers. Specifically, we show that for any sufficiently large positive integer numbers N and P, $\Omega(\sqrt{N}) = P = O(N^2)$, there exist weighted undirected graphs $G = ((V, E), \omega)$, and sets $\mathcal{P} \subseteq \binom{V}{2}$ of cardinality $|\mathcal{P}| = P$, such that any pair-wise preserver of the graph G with respect to the set $\mathcal{P}$ requires $|E| = \Omega((N \cdot P)^{2/3})$ edges. This lower bound is super-linear in $N + P$ for the range $\omega(\sqrt{N}) = P = o(N^2)$. We remark that in view of our linear upper bound of Thm. 1.1, this lower bound applies for the entire feasible range of P. In Section 3 we will proceed to showing similar, though weaker, lower bounds that are applicable for *unweighted undirected* graphs.

We next construct a fairly dense weighted graph $G = ((V, E), \omega)$, $\omega : E \to \mathbb{R}^+$, and a subset $\mathcal{P} \subseteq \binom{V}{2}$ of pairs of vertices. We will show that for any subgraph $G' = ((V, H), \omega)$, $H \subset E$ ($H \neq E$), there exists a pair $p = \{u, w\} \in \mathcal{P}$ so that $dist_{G'}(u, w) > dist_G(u, w)$.

Consider a square portion of the two-dimensional integer lattice $\mathbb{Z}^2$, with real dimensions $\sqrt{N} \times \sqrt{N}$, where N is the number of vertices in the graph G that we construct. In other words, let $V = \{(i, j) : i, j \in [(\sqrt{N})]\}$. (We assume that $\sqrt{N}$ is integral; all the non-integrality issues affect only lower-order terms of our results, and are, henceforth, ignored.)

Let T, $T \leq \sqrt{N}/10$, be a positive integer parameter of the construction that will be fixed later. For a pair of vertices $(i, j), (\ell, k) \in V$, $\{(i, j), (\ell, k)\}$ is an edge if the Euclidean distance $||(i, j) - (\ell, k)|| = \sqrt{(\ell - i)^2 + (k - j)^2}$ between (i, j) and (ℓ, k) is at most T, and $gcd(|\ell - i|, |k - j|) = 1$.

The weight function ω is defined by $\omega(e) = ||(i, j) - (\ell, k)||$ for every edge $e \in E$. This completes the construction of the graph $G = ((V, E), \omega)$.

The *boundary frame set* BF is defined by

$$\begin{aligned} BF &= \{(i, j) : (i \in [(T)]) \\ &\quad \text{or} \quad (i \in [(\sqrt{N})] \backslash [(\sqrt{N} - T)]) \\ &\quad \text{or} \quad (j \in [(T)]) \text{ or } (j \in [(\sqrt{N})] \backslash [(\sqrt{N} - T)])\}\ , \end{aligned}$$

and the *boundary set* B is defined by $B = \{(i, j) : (i = 0) \text{ or } (i = \sqrt{N} - 1) \text{ or } (j = 0) \text{ or } (j = \sqrt{N} - 1)\}$.

For a point $(i, j) \in BF$, and an edge $e = \{(i, j), (\ell, k)\}$ that is adjacent to (i, j) and such that $(\ell, k) \in V \backslash BF$, consider the (straight) line $\mathcal{L}$ that passes through the two points (i, j) and (ℓ, k) in the Euclidean plane. Let $L = \mathcal{L} \cap V$. Let L' be the subset of L that contains those points of L that are farther from (i, j) than from (ℓ, k) (in terms of the Euclidean distance). Let $(i', j') \in L'$ be the farthest (in terms of the Euclidean distance) point from (i, j). This point is called the *antipodal point to* (i, j) *in the direction of* (ℓ, k) (or, *in the direction of* $(\ell - i, k - j)$), and is denoted $(i', j') = A((i, j), (\ell, k))$. (The definitions of direction and antipodality apply for any pair of neighboring points (i, j) and (ℓ, k).) The set of pairs $\mathcal{P}$ is now defined by
$\mathcal{P} = \{\{(i, j), (i', j')\} : (i, j) \in BF, (i', j') = A((i, j), (\ell, k)), \{(i, j), (\ell, k)\} \in E\}$. Observe that every point $(i, j) \in V$ has $\Theta(T^2)$ neighbors, and thus $|E| = \Theta(N \cdot T^2)$. Also, since $|BF| = O(\sqrt{N} \cdot T)$, it follows that $P = O(\sqrt{N} \cdot T^3)$. The next lemma shows that all the antipodal points belong to the boundary frame BF.

LEMMA 2.1. *For an edge* $\{(\ell, k), (\ell', k')\} \in E$, $A((\ell, k), (\ell', k')) \in BF$.

Proof. Let $\mathcal{L}$ denote the line in the Euclidean plane that passes through the two points (ℓ,k) and (ℓ',k'), and consider the ray $\mathcal{R} \subseteq \mathcal{L}$ that starts in (ℓ',k') and does not contain the point (ℓ,k). Let S be the boundary of the square Sq whose corners are (in the counter-clockwise order) the origin, $(\sqrt{N}-1,0)$, $(\sqrt{N}-1,\sqrt{N}-1)$, $(0,\sqrt{N}-1)$, that is, the union of the segments that connect every consecutive pair of these corners, and the origin to $(0,\sqrt{N}-1)$. Let (x,y) be the intersection of the ray $\mathcal{R}$ with the boundary S of this square. If $\frac{x-\ell}{\ell'-\ell}$ is an integer, then, obviously, $(x,y) = A((\ell,k),(\ell',k'))$, and since $(x,y) \in BF$, we are done. Otherwise, let (x',y') be the closest to (x,y) point that belongs to the set L'.

We next show that $(x',y') \in BF$. Let I denote the segment of the ray $\mathcal{R}$ that connects the points (x,y) and $(x-(\ell'-\ell), y-(k'-k))$. Obviously, $(x',y') \in I$. Note, however, that the distance between (x',y') and the boundary B of the square Sq is at most $\max\{|\ell'-\ell|,|k'-k|\} \le T$. (Note that this is distance between a point and a line that is parallel to one of the axis.) It follows that $(x',y') \in BF$, as required. ∎

The next lemma follows directly from the construction of the set $\mathcal{P}$, and from Lemma 2.1.

LEMMA 2.2. *For an edge $\{(\ell,k),(\ell',k')\} \in E$, the pair $\{A((\ell,k),(\ell',k')), A((\ell',k'),(\ell,k))\}$ belongs to the set $\mathcal{P}$ of pairs.*

Proof. By Lemma 2.1, both points $A((\ell,k),(\ell',k'))$ and $A((\ell',k'),(\ell,k))$ belong to BF. Furthermore, it is easy to see that all the four points $A((\ell,k),(\ell',k'))$, (ℓ,k), (ℓ',k'), and $A((\ell',k'),(\ell,k))$, all belong to the same line $\mathcal{L}$. Consider the segment $\mathcal{J}$ of the line $\mathcal{L}$ that connects the points $A((\ell,k),(\ell',k'))$ and $A((\ell',k'),(\ell,k))$. Let $v \in \mathcal{J} \cap L$ be the (unique) point that belongs to the segment, and such that the edge $\{A((\ell',k'),(\ell,k)), v\}$ belongs to the edgeset E. Obviously, $A(A((\ell',k'),(\ell,k)),v) = A((\ell,k),(\ell',k'))$, and so the pair $\{A((\ell,k),(\ell',k')), A((\ell',k'),(\ell,k))\}$ belongs to $\mathcal{P}$. ∎

For a graph $G = ((V,E),\omega)$, and an edge $e \in E$, let $G \backslash e$ denote the graph $((V, E\backslash\{e\}), \omega|_{(E\backslash\{e\})})$.

The next lemma shows that no edge of the graph G can be removed without increasing the distance between some pair of points (i,j), (i',j') such that $\{(i,j),(i',j')\} \in \mathcal{P}$.

LEMMA 2.3. *For an edge $e = \{u,w\}$, $u = (\ell,k)$, $w = (\ell',k')$, let $x = A(u,w)$, $y = A(w,u)$. Then $dist_{G\backslash e}(x,y) > dist_G(x,y)$.*

The proof of this lemma is based on the observation that, intuitively, the shortest distance between x and y is attained (uniquely) by the straight line that connects them, and if this line is interrupted, the distance becomes longer.

COROLLARY 2.1. *For infinitely many positive integer numbers N and T, $T \le \sqrt{N}/10$, there exist weighted N-vertex graphs $G = ((V,E),\omega)$ with $|E| = \Theta(N \cdot T^2)$ edges, and a collection $\mathcal{P}$ with $P = O(\sqrt{N} \cdot T^3)$ pairs of vertices, such that every subgraph G' that preserves all the distances between the pairs of vertices from $\mathcal{P}$ contains all the edges of G.*

Note that the equations $|E| = \Theta(N \cdot T^2)$ and $P = O(\sqrt{N} \cdot T^3)$ imply the lower bound of $|E| = \Omega((NP)^{2/3})$, which is the main result of this section. Note that the lower bound is super-linear in $P+N$ whenever $\omega(\sqrt{N}) = P = o(N^2)$. For $P = O(\sqrt{N})$ no super-linear lower bound is possible due to our upper bound from Theorem 1.1.

This construction generalizes readily to any constant dimension $d = 3,4,\ldots$, but the obtained lower bounds are (weakly) inferior to that of $|E| = \Omega((N \cdot P)^{2/3})$.

3 Lower Bounds: Pair-wise Preservers for Unweighted Graphs

In this section we present a more elaborate construction that enables us to show lower bounds on the cardinalities of pair-wise preservers for *unweighted* undirected graphs. These lower bounds are somewhat weaker than the lower bounds for *weighted* graphs that are given by the equation $|E| = \Omega((N \cdot P)^{2/3})$, but they are also super-linear in $N+P$ in the *entire feasible range* of P, that is, $\omega(\sqrt{N}) = P = o(N^2)$.

For a fixed large positive integer T, and a fixed small positive integer $d = 2,3,\ldots$, consider the set $Ball_d(T) \cap \mathbb{Z}^d$, $Ball_d(T) = \{x \in \mathbb{R}^d : ||x|| \le T\}$, of all the points of the integer lattice $\mathbb{Z}^d$ that belong to the d-dimensional ball of radius T centered at the origin.

Consider the convex hull $CH = CH(Ball_d(T) \cap \mathbb{Z}^d)$ of the set $Ball_d(T) \cap \mathbb{Z}^d$. Obviously, this convex hull is symmetric around the origin, i.e., if $\alpha \in CH$ then $(-\alpha) \in CH$ as well. Let $VH = VH(Ball_d(T) \cap \mathbb{Z}^d)$ be the set of the *extreme points* (or *vertices*) of the convex hull CH. This set is also symmetric around the origin. It is known [7] that the cardinality of the set VH is $\Theta(T^{d-2+\frac{2}{d+1}})$, where the constant hidden by the Θ-notation depends only on the dimension d.

We next construct the unweighted graph $G = (V,E)$, and the set of pairs $\mathcal{P} \subseteq \binom{V}{2}$, that will be used for our lower bound. Let N be a fixed large positive integer. The vertex set V is the set of all points $x = (x_1,\ldots,x_d)$ of the integer lattice $\mathbb{Z}^d$ with non-negative coordinates and such that $||x||_\infty = \max\{x_i : i \in [d]\} \le N^{1/d} - 1$.

In other words, $V = \{(x_1, \ldots, x_d) \in \mathbb{Z}^d : 0 \le x_i \le N^{1/d} - 1, \ \forall i \in [d]\}$.

For every point $x = (x_1, \ldots, x_d) \in V$, let $\Gamma(x) = \{(x+y) \in V : y \in VH\}$ be the Minkowski sum of x and VH, where $T < \frac{1}{5}N^{1/d}$ is a positive integer parameter to be fixed later. Note that since the set VH is symmetric around the origin, $z \in \Gamma(x)$ if and only if $x \in \Gamma(z)$. The edgeset E is defined by $E = \{\{x, z\} : z \in \Gamma(x)\}$. Let the *boundary frame* set $BF \subseteq V$ be the set of all points $x = (x_1, \ldots, x_d) \in V$ so that at least one of the coordinates x_i, $i \in [d]$, satisfies either $0 \le x_i \le T - 1$ or $N^{1/d} - T \le x_i \le N^{1/d} - 1$, and the *boundary* set $B \subseteq V$ be the set of all the points $x \in V$ such that at least one of the coordinates is either 0 or $N^{1/d} - 1$.

For each point $x \in V$ and each neighbor $z \in \Gamma(x)$, we define $A(x, z)$ (the *antipodal point to x in the direction $z - x$* (or *in the direction z*)) in the following way. Let $\mathcal{L}$ be the line in the d-dimensional Euclidean space that passes through the points x and z. Let $L = \mathcal{L} \cap V$ be the set of the vertices of the graph G that belong to this line, and let $L' \subseteq L$ be the subset of L that contains only vertices that are closer to the point z than to the point x, in terms of the Euclidean distance. The point $A(x, z)$ is defined as the farthest (in terms of the Euclidean distance) point from x that belongs to the set L'. Finally, similarly to the way that the set $\mathcal{P}$ of pairs was constructed in Section 2, we now define $\mathcal{P} = \{\{x, A(x, z)\} : x \in BF, z \in \Gamma(x)\}$. Denote $P = |\mathcal{P}|$.

Note that for every vertex $x = (x_1, \ldots, x_d) \in V \backslash BF$, the vertex x has precisely $|VH|$ neighbors in the graph. In other words, the edgeset E has cardinality at least $|E| \ge |V \backslash BF| \cdot |VH| = \Omega((N^{1/d} - 2T)^d \cdot T^{d-2+\frac{2}{d+1}}) = \Omega(NT^{d-2+\frac{2}{d+1}})$. (Hereafter all the asymptotic notation may hide dependence on the dimension d, but not on the number of vertices N or the parameter T.)

Observe also that $|BF| = \Theta(N^{1-\frac{1}{d}}T)$, and $P = O(|BF| \cdot |VH|) = O(N^{1-\frac{1}{d}} \cdot T^{d-1+\frac{2}{d+1}})$.

We next argue that no edge can be removed from the graph G without increasing the distance between some pair of vertices $\{x, y\} \in \mathcal{P}$. To this end, consider an edge $e = \{u, w\} \in E$. Let $y = A(u, w)$, and $x = A(w, u)$. It is easy to see that the proof of Lemma 2.1 generalizes readily to the d-dimensional space with $d > 2$, and thus, $x, y \in BF$, and, furthermore, $\{x, y\} \in \mathcal{P}$.

The statement of the next lemma is analogous to Lemma 2.3. Its proof is, however, more involved, since the "Euclidean lengths" of the edges of the graph are no longer uniform.

Lemma 3.1. *$dist_{G \backslash e}(x, y) > dist_G(x, y)$.*

Proof. Let $\Pi_{x,y} = (\{x, x + (w-u)\}, \{x + (w-u), x + 2(w-u)\}, \ldots, \{x + (k-1)(w-u), x + k(w-u)\})$, $x + k(w-u) = y$, be the path in the graph G that connects the vertices x and y, and all its edges lie on the line $\mathcal{L}$ that passes through the points u and w. (We say that an edge $(z_1, z_2) \in E$ lies on a line $\mathcal{L}'$ if both points z_1 and z_2 belong to this line.)

Observe that the length of this path is $k = \frac{y_j - x_j}{w_j - u_j}$, where x_j, y_j, u_j and w_j are the jth coordinates of the vectors x, y, u and w, respectively, and $j \in [d]$ is one of the indices that satisfy $w_j - u_j \neq 0$. (Note that since the points x, y, u and w are co-linear, the expression $\frac{y_j - x_j}{w_j - u_j}$ is independent of the choice of the index $j \in [d]$, as far as $w_j - u_j \neq 0$. Since $w - u \in VH$, it follows that $w - u \neq 0$.)

Consider some path $\Pi'_{x,y}$ of length at most k that connects the vertices x and y in the graph G. Let $\Pi'_{x,y} = (e^{(1)}, e^{(2)}, \ldots, e^{(\ell-1)})$, $\ell \le k + 1$, $e_i = \{x^{(i)}, x^{(i+1)}\}$, $i \in [\ell - 1]$, $x^{(1)} = x$, $x^{(\ell)} = y$.

For an index $i \in [\ell - 1]$, let $d^{(i)} = x^{(i+1)} - x^{(i)}$, where $e^{(i)} = \{x^{(i)}, x^{(i+1)}\}$. It follows that $y = x^{(\ell)} = x + \sum_{i=1}^{\ell-1} d^{(i)}$. Hence $y - x = \sum_{i=1}^{\ell-1} d^{(i)}$, $\ell - 1 \le k$. Note also that $y - x = k \cdot (w - u)$. Hence

$$w - u = \frac{1}{k} \sum_{i=1}^{\ell-1} d^{(i)} .$$

Let $\{d'^{(1)}, \ldots, d'^{(\ell')}\}$, $\ell' \le \ell - 1$, be the set of distinct terms that appear in the sum $\sum_{i=1}^{\ell-1} d^{(i)}$, and let α_i, $i \in [\ell']$, denote the number of times that the element $d'^{(i)}$ appears in this sum. It follows that $w - u = \frac{1}{k} \sum_{i=1}^{\ell'-1} \alpha_i d'^{(i)}$, and $\sum_{i=1}^{\ell'} \alpha_i = \ell - 1 \le k$. Hence

$$w - u \;=\; \sum_{i=1}^{\ell'} \beta_i d'^{(i)} , \tag{3.1}$$

where $\beta_i = \alpha_i / k > 0$, $\sum_{i=1}^{\ell'} \beta_i = \frac{1}{k} \sum_{i=1}^{\ell'} \alpha_i = \frac{\ell-1}{k} \le 1$.

Recall that $\{u, w\}$ is an edge of the graph G, and so the vector $w - u$ belongs to the set VH. Analogously, for each index $i \in [\ell']$, there exists an index $j \in [\ell - 1]$ such that $d'^{(i)} = d^{(j)}$, and $d^{(j)} = x^{(j+1)} - x^{(j)}$, where $e^{(j)} = \{x^{(j)}, x^{(j+1)}\}$ is an edge of the graph G. Hence the vector $d'^{(i)} = d^{(j)} = x^{(j+1)} - x^{(j)}$ also belongs to the set VH. In other words, $w - u, d'^{(1)}, d'^{(2)}, \ldots, d'^{(\ell')} \in VH$, and also Equation (3.1) is satisfied. However, since VH is the set of *extreme points* of the convex hull of a set of points in $\mathbb{R}^d$, it follows that a vector $w - u \in VH$ can be represented uniquely as a convex combination of vectors from VH, specifically, as $w - u = \beta_1(w - u)$ with $\beta_1 = 1$. It follows that $\ell' = 1$, and $d'_1 = w - u$, and so the path $\Pi'_{x,y}$ coincides with the path $\Pi_{x,y}$.

Hence the path $\Pi_{x,y}$ is the unique shortest path between the vertices x and y in the graph G, and this path uses the edge e. Hence $dist_{G\setminus e}(x,y) > dist_G(x,y)$. ∎

Recall that $|E| = \Omega(N \cdot T^{d-2+\frac{2}{d+1}})$ and $P = O(N^{1-\frac{1}{d}} \cdot T^{d-1+\frac{2}{d+1}})$, for $T \le \frac{1}{5}N^{1/d}$. A straightforward calculation shows that

$$|E| = \Omega\left(N^{\frac{2d}{d^2+1}} \cdot P^{\frac{d(d-1)}{d^2+1}}\right) , \tag{3.2}$$

and this lower bound is applicable to $\Omega(N^{1-\frac{1}{d}}) = P = O(N^{2-\frac{2}{d+1}})$. For $d = 2,3,\ldots$, let E_d denote the right-hand side of Equation (3.2). By comparing these lower bounds for different values of d it follows that for $d = 2,3,\ldots$, in the range $\Omega(N^{2\cdot\frac{d^2-d-1}{(d-1)(d+2)}}) = P = O(N^{2\cdot\frac{d^2+d-1}{d(d+3)}})$, the lower bound is

$$E = \Omega(E_d) = \Omega\left(N^{\frac{2d}{d^2+1}} \cdot P^{\frac{d(d-1)}{d^2+1}}\right) . \tag{3.3}$$

4 Lower Bounds: Source-wise Preservers for Unweighted Graphs

In this section we show lower bounds on the cardinalities of source-wise preservers.

4.1 Constructing a Large Convexly Independent Set We start with describing our variant of Jarnik construction [19] of a large convexly independent set (henceforth, CIS) of two-dimensional vectors of norm at most R, for some fixed parameter R. We will use this construction, and some of its properties, for our lower bound.

Let t be an even integer parameter to be fixed later. Let $Z = \{(a,b) : a,b \in [t]\setminus[t/2], gcd(a,b) = 1\}$. Note that $|Z| = \Theta(t^2)$. (For a sufficiently large t, Z is very close to $\frac{t^2}{4} \cdot \frac{6}{\pi^2}$, since $\prod_p(1-1/p^2) = 6/\pi^2$ (the product is taken over all prime numbers). See, e.g., [17].) Sort all the elements of Z by the ratio b/a, starting from the pair (a,b) with the largest ratio b/a, and ending with the smallest one. (Note that if two ratios b_1/a_1, b_2/a_2 are equal then the two vectors (a_1,b_1), (a_2,b_2) are co-linear, and since both of them are integer vectors, one of them is a rational multiple of the other. However, since both have gcd equal to 1, it follows that the two vectors are equal, contradiction.)

Let $(a_1,b_1),\ldots,(a_k,b_k)$ be the sorted sequence of the vectors of Z. Let $A = \sum_{i=1}^k a_i$, $B = \sum_{i=1}^k b_i$. Let $w_0 = (A,0)$, $w_1 = w_0 + (-a_1,b_1)$, $w_2 = w_1 + (-a_2,b_2),\ldots,w_k = w_{k-1} + (-a_k,b_k) = (A - \sum_{i=1}^k a_i, \sum_{i=1}^k b_i) = (0,B)$. Generally, for $j = 0,1,\ldots,k$, $w_j = (A - \sum_{i=1}^j a_i, \sum_{i=1}^j b_i) = (\sum_{i=j+1}^k a_i, \sum_{i=1}^j b_i)$, where $\sum_{i=k+1}^k(\cdot)$ is defined as 0. Denote $W = \{w_0,w_1,\ldots,w_k\}$. Note that $|W| = k = |Z| = \Theta(t^2)$. The norms of the vectors w_j, $j = 0,1,\ldots,k$, are at most $\sqrt{2}\cdot k \cdot t = O(t^3)$. It is also not hard to see (see [19] for the formal proof) that the set W is a CIS, or, in other words, that the vectors $w_0,w_1,\ldots,w_k$ are the extreme points of the convex hull of the set $\{w_0,w_1,\ldots,w_k\}$, $CH(w_0,w_1,\ldots,w_k)$.

Furthermore, consider the set U of vectors given by $U = \{u = (x,y) : \exists\sigma_x,\sigma_y \in \{-1,1\} \text{ s.t. } (\sigma_x x, \sigma_y y) \in W\}$.

It is easy to see that this set is a CIS as well. Consider the convex hull $CH(U)$ of the set U. Let $(v_0,v_1,\ldots,v_{4k})$ be the sequence of vectors of U ordered counter-clock-wise, starting with $v_0 = w_0$. By construction, every edge of the convex hull $CH(U)$ has slope a/b, where $t/2 \le |a|,|b| \le t$. (The *slope* of the line $\mathcal{L} = \{(x_1,y_1) + \alpha\cdot(x_2,y_2) : \alpha \in \mathbb{R}\}$ is defined as x_2/y_2 if $y_2 \ne 0$, and as ∞ otherwise. In this paper, however, all the lines have finite slopes.) Setting $R = t^3$ we derive the following theorem.

THEOREM 4.1. *[19] For infinitely many positive integer numbers R there exist CISs of vectors $(v_0,v_1,\ldots,v_{4k}) \in \mathbb{R}^2$, ordered counter-clock-wise, with $k = \Theta(R^{2/3})$ of norm $||v_j|| \le R$ for every $j \in \{0,1,\ldots,4k\}$. Furthermore, for every index j in this range, $||v_{j+1} - v_j||_\infty = \Theta(R^{1/3})$, where $j+1$ is shorthand for $j+1$* (mod $4k+1$).

4.2 Constructing Supporting Lines We next describe the construction of a collection of supporting lines of the polygon $CH(U)$.

Fix a positive integer parameter T. Let C be the CIS that satisfies the properties guaranteed by Theorem 4.1 with $R = T$. Let P be the boundary of the convex hull $\hat{P} = CH(C)$. Note that P is a convex polygon whose extreme points are the vectors of C.

We need an additional piece of notation. For a convex polygon Q, let $Ex(Q)$ denote the set of its extreme points. Note that $Ex(P) = C$.

For each vertex v of the polygon P, let $v^{(1)}$ and $v^{(2)}$ be the two vertices of P so that there are two facets of the polygon that connect $v^{(1)}$ to v, and v to $v^{(2)}$. Let $a^{(1)}/b^{(1)}$ and $a^{(2)}/b^{(2)}$ be the two slopes of these two facets (in their lowest terms), and assume without loss of generality that $a^{(2)}/b^{(2)} > a^{(1)}/b^{(1)}$. (Note that a facet of a polygon is a segment, and so its slope is well-defined.) Then, by construction, $v^{(2)} - v = (-a^{(2)},b^{(2)})$, $v - v^{(1)} = (-a^{(1)},b^{(1)})$, and $|a^{(i)}|,|b^{(i)}| \le T^{1/3}$ for $i = 1,2$. Furthermore, there exists no vector $(a^{(3)},b^{(3)}) \in \mathbb{Z}^2$ with $(a^{(3)})^2 + (b^{(3)})^2 \le T^{2/3}$, and $gcd(a^{(3)},b^{(3)}) = 1$ so that $a^{(1)}/b^{(1)} < a^{(3)}/b^{(3)} < a^{(2)}/b^{(2)}$. It follows that the vector $(a^{(1)}+a^{(2)},b^{(1)}+b^{(2)})$ has an intermediate slope

$a^{(1)}/b^{(1)} < (a^{(1)} + a^{(2)})/(b^{(1)} + b^{(2)}) < a^{(2)}/b^{(2)}$, and that the numbers $a^{(1)} + a^{(2)}$ and $b^{(1)} + b^{(2)}$ are relatively prime. Furthermore, $||(a^{(1)} + a^{(2)}, b^{(1)} + b^{(2)})|| = \sqrt{(a^{(1)} + a^{(2)})^2 + (b^{(1)} + b^{(2)})^2} = O(T^{1/3})$.

For each vector $v \in C$, let $e(v)$ denote the vector $(a^{(1)} + a^{(2)}, b^{(1)} + b^{(2)})$.

Definition 4.1. *For a set $Z \subseteq \mathbb{R}^2$ of points, the line $\mathcal{L} = \{v + \alpha \cdot u : \alpha \in \mathbb{R}\}$, $u = (a, b) \in \mathbb{R}^2$, $v \in Z$, is called a* supporting line of the set Z in the point $v \in Z$ *if the two following conditions hold. (1) The line passes through the point v. (2) The sign of the inner product $\langle x - v, u^{\perp}\rangle$ is negative for every point $z \in Z \backslash \{v\}$, where $u^{\perp} = (-b, a)$.*

Lemma 4.1. *Let $\mathcal{L}$ be the line (in the the Euclidean plane) that passes through the point v and is parallel to the vector $e(v)$. Then the line $\mathcal{L}$ is a supporting line of the polygon $\hat{P}$ in the point v.*

4.3 Constructing the Instance We next describe the construction of the graph $G = (V, E)$, and a subset $\mathcal{S} \subseteq V$ of sources that will be used in the proof of our lower bound for the sources problem.

Let $N > 10 \cdot T^2$ be another positive integer parameter. Enlarge the polygon P by a factor $\sqrt{N}/T$. In other words, let $P' = (\sqrt{N}/T) \cdot P = \{(\sqrt{N}/T) \cdot u : u \in P\}$. Note that the polygon P' is contained in a disc of radius $\sqrt{N}$, and therefore, it contains $O(N)$ points of the integer lattice $\mathbb{Z}^2$.

For each extreme point v of the polygon P, consider the line that passes through the origin and is perpendicular to the vector $e(v)$. Let A_v be one of the two points on this line (chosen arbitrarily) that are at (Euclidean) distance exactly $\sqrt{N}/2$ from the origin. Let I_v be a segment of length $\sqrt{N}$ parallel to the vector $e(v)$ that passes through the point A_v, and, furthermore, A_v is its center. Let C_v and D_v be its endpoints. Let I'_v be a segment of the same length as I_v (that is, $\sqrt{N}$), parallel to I_v, and such that its center A'_v is located on the line that connects the origin to A_v, at distance exactly T from the point A_v. Let C'_v, D'_v be its endpoints. Let B_v be the rectangle $C_v D_v D'_v C'_v$.

Let $F = \frac{\sqrt{N}}{T}$, and let B'_v be the box B_v displaced by $F \cdot v$. In other words, $B'_v = \{x + F \cdot v : x \in B_v\}$. Note that by triangle inequality, the box B'_v is contained in a disc of radius $2\sqrt{N}$.

For each integer point $x \in B_v \cap \mathbb{Z}^2$, insert the edges $\{x, x + v\}, \{x + v, x + 2v\}, \ldots, \{x + (F - 1) \cdot v, x + F \cdot v\}$ into the edgeset E of the graph $G = (V, E)$ that we construct.

Definition 4.2. *For a box B_v, a segment I_v is called an* aligned segment *of the box B_v if it contains at least one integer point x of the box B_v, it is parallel to the vector $e(v)$ (and, consequently, to the long edge of the box B_v), and it contains all the integer points of the box B_v of the form $x + \alpha \cdot e(v)$, $\alpha \in \mathbb{Z}$, and its endpoints lie on the boundary of the box B_v.*

For an aligned segment I_v , let Z_v denote the set of integer points of the segment I_v. The set Z_v will be referred also as an integer aligned segment of the box *B_v.*

Fix the vector v for the rest of this section. Note that the real length of each aligned segment I_v is $\sqrt{N}$. Note also that the set of all the integer points of the box B_v decomposes into the disjoint union of integer aligned segments Z_v of B_v. For an integer aligned segment Z_v of the box B_v, let $Z'_v = Z_v + F \cdot v = \{x + F \cdot v : z \in Z_v\}$. Note that the set Z'_v is an integer aligned segment of the box B'_v.

Let $m(v) = |Z_v|$ denote the cardinality of the set Z_v. Note that as the notation suggests, this cardinality does not depend on the choice of the particular integer aligned segment of the box B_v. As v is fixed, let m serve as a shortcut for $m(v)$.

For each integer aligned segment Z_v of B_v, order the points of Z_v according to the order in which they appear on the line (there are two such orders; choose either one of them arbitrarily). Let o_v be this ordering, and write $Z_v = (x_1, \ldots, x_m)$ with $x_i <_{o_v} x_j$ if and only if $i < j$. Order Z'_v in the same order, i.e., $Z'_v = (x'_1, \ldots, x'_m)$, $x'_i = x_i + F \cdot v$, $i \in [m]$.

Arrange the numbers $(1, 2, \ldots, m)$ as a $\sqrt{m} \times \sqrt{m}$ matrix M (ignoring possible non-integrality of $\sqrt{m}$; it affects only the lower-order terms of the analysis), with $M_{ij} = \sqrt{m}(i - 1) + j$, for $i, j \in [\sqrt{m}]$.

Let R_i (resp., C_i), $i \in [\sqrt{m}]$ be the set of numbers that appear on the ith row (resp., column) of this matrix, i.e., $R_i = \{\sqrt{m}(i - 1) + 1, \sqrt{m}(i - 1) + 2, \ldots, (\sqrt{m})i\}$ (resp., $C_i = \{i, \sqrt{m} + i, \ldots, (\sqrt{m} - 1)\sqrt{m} + i\}$). Let $R_i(Z_v) = \{x_\ell \in Z_v : \ell \in R_i\}$, and, analogously, $C_i(Z_v) = \{x_\ell \in Z_v : \ell \in C_i\}$. Note that for every pair of indices $i, j \in [\sqrt{m}]$, $|R_i \cap C_j| = 1$, and consequently, $|R_i(Z_v) \cap C_j(Z_v)| = 1$. For each index $i \in [\sqrt{m}]$, a new *source* $s_i(Z_v)$ is introduced, that is, a vertex of the graph G that belongs to the set $\mathcal{S}$ of sources. The source vertices do not correspond to points of the Euclidean plane. In addition, $\sqrt{m} - 2$ (assuming that $\sqrt{m}$ is a power of 2) new auxilary vertices are introduced for each source. (If $\sqrt{m}$ is not a power of 2, then $2^t - 2$ vertices are introduced, where t satisfies $2^{t-1} < \sqrt{m} \leq 2^t$.) These auxilary vertices are used to form a complete binary tree rooted in $s_i(Z_v)$. This tree $\hat{\tau}$ has $\sqrt{m}/2$ leaves, all of which are new auxilary vertices. Each of these leaves is connected to exactly two vertices of the set $R_i(Z_v)$, and each vertex of $R_i(Z_v)$ has exactly one neighbor in $\hat{\tau}$.

Consider the obtained complete binary tree $\hat{\tau}_i(Z_v)$, rooted in the source $s_i(Z_v)$. This tree has depth $\log \sqrt{m}$, and its leaves are the $\sqrt{m}$ vertices of the set $R_i(Z_v)$. The vertices of this tree are assigned *levels* in the following way. Each leaf z is assigned level $\ell(z) = 0$, and each vertex w whose all children are assigned level ℓ is assigned level $\ell(w) = \ell+1$. An edge $e = \{u, w\} \in \hat{\tau}_i(Z_v)$ that connects a vertex u to its parent w in the tree is assigned the level of the child w, i.e., $\ell(e) = \ell(u)$. Next, each edge e of the tree $\hat{\tau}_i(Z_v)$ is replaced by a path of length $20 \cdot 2^{\ell(e)} \cdot \frac{\sqrt{m}}{T^{2/3}}$, which is formed using new auxilary vertices. The resulting tree is denoted $\tau_i(Z_v)$, and the tree $\hat{\tau}_i(Z_v)$ is called its *skeleton*.

An almost symmetrical tree is constructed for the vertices of Z'_v. Specifically, a new source $s'_i(Z'_v)$ is introduced, along with a bunch of new auxilary vertices. These vertices are used to build a complete binary tree rooted in the source $s'_i(Z'_v)$. This tree has $\sqrt{m}/2$ leaves, and each of them is connected to two vertices of the set $C_i(Z'_v)$ such that each vertex of the set $C_i(Z'_v)$ is connected to exactly one such leaf. Finally, every edge e of the obtained tree $\hat{\tau}'_i(Z'_v)$ by a path of appropriate length (specifically, $20 \cdot 2^{\ell(e)} \cdot \frac{\sqrt{m}}{T^{2/3}}$, where $\ell(e)$ is the level of the edge), and this path is formed using new auxilary vertices. The obtained tree is denoted $\tau'_i(Z'_v)$.

This is done separately for every extreme point $v \in Ex(P)$, for every integer aligned segment Z_v of B_v, and Z'_v of B'_v, and for every index $i \in [\sqrt{m(v)}]$.

4.4 The Analysis of the Construction Note that for each vector v, each integer aligned segment Z_v, and each pair of sources $s_i(Z_v)$, $s'_j(Z'_v)$, $i, j \in [\sqrt{m(v)}]$, there exists a unique vertex $x \in Z_v$ such that x serves as a leaf of the tree $\tau_i(Z_v)$, and $x' = x + F \cdot v$ serves as a leaf of the tree $\tau'_j(Z'_v)$. Furthermore, for each vertex $x \in Z_v$, there exists a unique pair of sources $s_i(Z_v)$, $s'_j(Z'_v)$, as above. Therefore, for each vertex $x \in Z_v$, the shortest path between x and x' is indispensable, i.e., no edge of this path can be removed without increasing the distance between some pair of sources.

The involved way in which the trees $\tau_i(Z_v)$, $\tau'_j(Z'_v)$ are constructed is dictated by the necessity to balance two contradictory requirements. On the one hand we need to prevent these trees from affecting the geometric properties of the graph, that is, from making a distance in the graph between some pair of points of the Euclidean plane shorter than the Euclidean distance between them (divided by a scaling factor). On the other hand, we have to keep the number of vertices as small as possible, in order to achieve a stronger lower bound.

We next provide a few bounds on the parameters of the construction.

Lemma 4.2. *The number N' of vertices in the graph $G = (V, E)$ is $O(N + N^{3/4} \cdot T^{5/6} \cdot \log N)$, the number of edges is $|E| = O(N \cdot T^{2/3} + N^{3/4} \cdot T^{5/6} \cdot \log N)$, and the number of sources is $S = O(N^{1/4} \cdot T^{11/6})$.*

Intuitively, the next lemma shows that the auxilary trees do not affect the geometric properties of the graph.

Lemma 4.3. *For an extreme point $v \in Ex(P)$, and a pair of points u, w that belong to the same integer aligned segment Z_v of the box B_v, any path Π between them that is contained entirely in some auxilary tree $\tau_i(Z_v)$ can be replaced by a sequence $(u = u_0, u_1, \ldots, u_L = w)$, $L = |\Pi|$, of points in the plane that satisfy that for every index $j \in [L]$, the vector $u_j - u_{j-1}$ is contained in the convex hull $\hat{P}$ of C.*

Obviously, the same lemma applies to the trees $\tau'_j(Z'_v)$. Furthermore, this lemma readily generalizes for a pair of integer points u, w in the plane that belong to the vertex set V of the graph, but do not necessarily belong to the same aligned segment.

The proof of the next lemma was outlined in the beginning of this section.

Lemma 4.4. *No edge e of the graph $G = (V, E)$ can be removed without increasing the distance between at least one pair of sources $s, s' \in \mathcal{S}$.*

To summarize, we have proved that there exist infinitely many values of positive integer parameters N and T for which there exists a graph $G = (V, E)$ with $N' = O(N + N^{3/4} \cdot T^{5/6} \cdot \log N)$ vertices, $|E| = \Theta(N \cdot T^{2/3} + N^{3/4} \cdot T^{5/6} \cdot \log N)$ edges, and a subset $\mathcal{S} \subseteq V$ of $O(N^{1/4} \cdot T^{11/6})$ sources, so that removal of any edge e from the graph results in increasing the distance between some pair of sources.

Direct calculation shows that for $T \leq \frac{N^{3/10}}{\log^{6/5} N}$, $N' = O(N)$, and the lower bound of $|E| = \Omega(N^{10/11} \cdot S^{4/11})$ follows. This lower bound is super-linear in $N + S^2$ for $\omega(N^{1/4}) = S = o(N^{5/9})$. We proved the following theorem.

Theorem 4.2. *There exists infinitely many positive integer numbers N and S, $\Omega(N^{1/4}) = S = o(N^{5/9})$, for which there exist N-vertex unweighted undirected graphs $G = (V, E)$ and subsets $\mathcal{S} \subseteq V$ of S sources, with $|E| = \Omega(N^{10/11} \cdot S^{4/11})$ edges such that removal of any edge e from the graph results in increasing the distance between some pair of sources.*

We note that a three-dimensional construction can be used to generalize and improve this lower bound, and specifically, we show that $|E| = \Omega(N^{9/11} \cdot S^{6/11})$.

This improved lower bound is super-linear in the range $\omega(N^{1/4}) = S = o(N^{9/16})$, and it is stronger than the lower bound of Theorem 4.2 for $\omega(\sqrt{N}) = S = o(N^{9/16})$. Furthermore, a yet stronger lower bound of $|E| = \Omega(N^{6/7} \cdot S^{4/7})$ can be proved. That lower bound, however, applies only to *weighted* graphs. This lower bound is super-linear in a wider range of S, specifically, $\omega(N^{1/4}) = S = o(N^{3/5})$. This lower bound is also stronger than the lower bound for unweighted graphs for every value of S for which either of the lower bounds is super-linear in $N + S^2$.

Acknowledgements

The authors are grateful to anonymous referees for helpful remarks.

References

[1] N. Alon, R. Karp, D. Peleg, and D. West. A graph-theoretic game and its application to the k-server problem. *SIAM Journal on Computing*, 24:78–100, 1995.

[2] B. Awerbuch, B. Berger, L. Cowen, and D. Peleg. Near-linear time construction of sparse neighborhood covers. *SIAM Journal on Computing*, 28(1):263–277, 1998.

[3] B. Awerbuch and D. Peleg. Network synchronization with polylogarithmic overhead. In *Proc. 31st Symp. on Foundations of Computer Science*, pages 514–522, 1990.

[4] B. Awerbuch and D. Peleg. Routing with polynomial communication-space trade-off. *SIAM J. Discrete Math.*, (5):151–162, 1992.

[5] B. Awerbuch, D. Peleg, B. Patt-Shamir, and M. Saks. Adapting to asynchronous dynamic networks with polylogarithmic overhead. In *Proc. 24th ACM Symp. on Theory of Computing*, pages 557–570, 1992.

[6] A. Balog and I. Barany. On the convex hull of the integer points in a disc. *Discrete and Computational Geometry*, 6:39–44, 1991.

[7] I. Barany and D. G. Larman. The convex hull of the integer points in a large ball. *Mathematische Annalen*, (312):167–181, 1998.

[8] S. Baswana, T. Kavitha, K. Mehlhorn, and S. Pettie. New constructions of (a,b)-spanners and additive spanners. In *SODA: ACM-SIAM Symposium on Discrete Algorithms*, 2005.

[9] S. Baswana and S. Sen. A simple linear time algorithm for computing a $(2k-1)$-spanner of $o(n^{1+1/k})$ size in weighted graphs. In *Proc. of the 30th International Colloq. on Automata, Languages and Computing, ICALP 2003*, pages 284–296, 2003.

[10] B. Bollobas, D. Coppersmith, and M. Elkin. Sparse distance preservers and additive spanners. In *Proc. of the 14th ACM-SIAM Symp. on Discrete algorithms*, pages 414–423, 2003.

[11] J. Bourgain. On Lipschitz embedding of finite metric space in Hilbert space. *Israel J. Math,*, 56:46–52, 1985.

[12] E. Cohen. Fast algorithms for constructing t-spanners and paths of stretch t. In *Proc. 34th IEEE Symp. on Foundations of Computer Science*, pages 648–658, 1993.

[13] D. Dor, D. Halperin, and U. Zwick. All pairs almost shortest paths. In *Proc. 37th IEEE Symp. on Foundations of Computer Science*, pages 452–461, 1997.

[14] M. Elkin. Computing almost shortest paths. In *Proc. 20th ACM Symp. on Principles of Distributed Computing*, pages 53–63, 2001.

[15] M. Elkin and D. Peleg. $(1+\epsilon,\beta)$-spanner constructions for general graphs. In *Proc. of the 33th ACM Symp. on Theory of Computing*, pages 173–182, 2001.

[16] C. Gavoille, D. Peleg, S. Perennes, and R. Raz. Distance labeling in graphs. In *Proc. 12th ACM-SIAM Symp. on Discrete algorithms*, pages 210–219, 2001.

[17] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, 1979.

[18] P. Indyk and J. Matousek. *Low Distortion Embeddings of Finite Metric Spaces: Chapter 8 in CRC Handbook of Discrete and Computational Geometry*. 2004.

[19] V. Jarnik. Uber Gitterpunkte und konvex Kurven. *Math. Z.*, 2:500–518, 1925.

[20] W. Johnson and J. Lindenstrauss. Extensions of Lipschitz embeddings into a Hilbert space. *Contempt. Math.*, 26:189–206, 1984.

[21] Y. Mansour and D. Peleg. An approximation algorithm for minimum-cost network design. *The Weizmann Institute of Science Technical Report, CS94-22*, 1994.

[22] J. Matousek. *Lectures in Discrete Geometry: Graduate Texts in Mathematics, Springer, New York*, volume 212. 2002.

[23] D. Peleg. Proximity-preserving labeling schemes. *Journal of Graph Theory*, (33):167–176, 2000.

[24] D. Peleg and A. Schäffer. Graph spanners. *Journal of Graph Theory*, 13:99–116, 1989.

[25] D. Peleg and E. Upfal. A trade-off between space and efficiency for routing tables. *Journal on ACM*, 36(3):510–530, 1989.

[26] L. Roditty, M. Thorup, and U. Zwick. Roundtrip spanners and roundtrip routing in directed graphs. In *Proc. 13th ACM-SIAM Symposium on Discrete Algorithms*, 2002.

[27] M. Thorup and U. Zwick. Approximate distance oracles. In *Proc. of the 33th ACM Symp. on Theory of Computing*, pages 183–192, 2001.