# On Notions of Distortion and an Almost Minimum Spanning Tree with Constant Average Distortion

Yair Bartal[1], Arnold Filtser[2], and Ofer Neiman[2]

[1] School of Engineering and Computer Science, Hebrew University. Email: yair@cs.huji.ac.il
[2] Department of Computer Science, Ben-Gurion University of the Negev. Email: {arnoldf,neimano}@cs.bgu.ac.il

**Abstract**

Minimum Spanning Trees of weighted graphs are fundamental objects in numerous applications. In particular in distributed networks, the minimum spanning tree of the network is often used to route messages between network nodes. Unfortunately, while being most efficient in the total cost of connecting all nodes, minimum spanning trees fail miserably in the desired property of approximately preserving distances between pairs. While known lower bounds exclude the possibility of the worst case distortion of a tree being small, it was shown in [4] that there exists a spanning tree with constant average distortion. Yet, the weight of such a tree may be significantly larger than that of the MST. In this paper, we show that any weighted undirected graph admits a *spanning tree* whose weight is at most $(1+\rho)$ times that of the MST, providing *constant average distortion* $O(1/\rho)$.

The constant average distortion bound is implied by a stronger property of *scaling distortion*, i.e., improved distortion for smaller fractions of the pairs. The result is achieved by first showing the existence of a low weight *spanner* with small *prioritized distortion*, a property allowing to prioritize the nodes whose associated distortions will be improved. We show that prioritized distortion is essentially equivalent to coarse scaling distortion via a general transformation, which has further implications and may be of independent interest. In particular, we obtain an embedding for arbitrary metrics into Euclidean space with optimal prioritized distortion.2

## 1 Introduction

One of the fundamental problems in graph theory is that of constructing a Minimum Spanning Tree (MST) of a given weighted graph $G = (V, E)$. This problem and its variants received much attention, and has found numerous applications. In many of these applications, one may desire not only minimizing the weight of the spanning tree, but also other desirable properties, at the price of losing a small factor in the weight of the tree compared to that of the MST. Define the *lightness* of $T$ to be the total weight of $T$ (the sum of its edge weights) divided by the weight of an MST. One well known example is that of a Shallow Light Tree (SLT) [21, 8], which is a rooted spanning tree having near optimal $(1+\rho)$ lightness, while approximately preserving all distances from the root to the other vertices.

It is natural to ask that the spanning tree will preserve well all pairwise distances in the graph. However, it is easy to see that no spanning tree can maintain such a requirement. In particular, even in the case of the unweighted cycle graph on $n$ vertices, for every spanning tree there is a pair of neighboring vertices whose distance increases by a factor of $n-1$. A natural relaxation of this demand is that the spanning tree approximates all pairwise distances *on average*. Formally, the distortion of the pair $u, v \in V$ in $T$ is defined as $\frac{d_T(u,v)}{d_G(u,v)}$, and the *average distortion* is $\frac{1}{\binom{n}{2}} \sum_{\{u,v\} \in \binom{V}{2}} \frac{d_T(u,v)}{d_G(u,v)}$, where $d_G$ (respectively $d_T$) is the shortest-path metric in $G$ (resp. $T$).[1] In [4], it was shown that for every weighted graph, it is possible to find a spanning tree which has constant average distortion.

In this paper, we devise a spanning tree of optimal $(1+\rho)$ lightness that has $O(1/\rho)$ average distortion over all pairwise distances. We show that this result is tight by exhibiting a lower bound on the tradeoff between lightness and average distortion, that in order to get $1+\rho$ lightness the average distortion must be $\Omega(1/\rho)$ (this holds even if the spanning subgraph is not necessarily a tree), and in particular, the average distortion for an MST is as bad as $\Omega(n)$.

Our main result may be of interest for network applications. It is extremely common in the area of distributed computing that an MST is used for communication between the network nodes. This allows easy centralization of computing processes and an efficient way of broadcasting through the network, allowing communication to all nodes at a minimum cost. Yet, as already mentioned above, when communication is required between specific pairs of nodes, the cost of routing through the MST may be extremely high, even when their real distance is small. However, in practice it is the average distortion, rather than the worst-case distortion, that is often used as a practical measure of quality, as has been a major motivation behind the initial work of [22, 3, 4]. As noted above, the MST still fails even in this relaxed measure. Our result overcomes this by promising small routing cost between nodes on average, while still possessing the low cost of broadcasting through the tree, thereby maintaining the standard advantages of the MST.

Our main result on a low average distortion embedding follows from analyzing the *scaling distortion* of the embedding. This notion, first introduced in [22][2], requires that for *every* $0 < \epsilon < 1$, the distortion of all but an $\epsilon$-fraction of the pairs is bounded by the appropriate function of $\epsilon$. In [3] it was shown that one may obtain bounds on the average distortion, as well as on higher moments of the distortion function, from bounds on the scaling distortion. Our scaling distortion bound for the constructed spanning tree is[3] $\tilde{O}(1/\sqrt{\epsilon})/\rho$, which is nearly tight as a function of $\epsilon$ [4].

We also obtain a probabilistic embedding devising a distribution over (light) spanning trees with $\text{polylog}(1/\epsilon)/\rho$ scaling distortion, thus providing constant bounds on all fixed moments of the distortion (i.e., the $l_q$-distortion [3] for fixed $q$).

Our main technical contribution, en route to this result, may be of its own interest: We devise a *spanner* (a subgraph of $G$) with $1+\rho$ lightness and low *prioritized distortion*. This notion, introduced recently in [16], means that for every given ranking $v_1, \ldots, v_n$ of the vertices of the graph, there is an embedding where the distortion of pairs including $v_j$ is bounded as a function of the rank $j$. Here we show a light spanner construction with prioritized distortion at most $\tilde{O}(\log j)/\rho$. We then show a connection between the notions of prioritized distortion and scaling distortion (discussed further below), and use this to argue that our spanner has scaling distortion $\tilde{O}(\log(1/\epsilon))/\rho$, and thus average distortion $O(1/\rho)$. Although we do not obtain a spanning tree here, this result has a few advantages, as we get constant bounds on all fixed moments of the distortion function (also called the $\ell_q$-distortion). Moreover, the worst-case distortion is only logarithmic in $n$. We note that

[1] Distortion is sometimes referred to as stretch.

[2] Originally coined gracefully degrading embedding.

[3] By $\tilde{O}(f(n))$ we mean $O(f(n) \cdot \text{polylog}(f(n)))$.

all of our results admit deterministic polynomial time algorithms.

Another technical contribution is a general, black-box reduction, that transform constructions of spanners with distortion $t$ and lightness $\ell$ into spanners with distortion $t/\delta$ and lightness $1+\delta\ell$ (here $0<\delta<1$). This reduction can be applied in numerous settings, and also for many different special families of graphs. In particular, this reduction allows us to construct prioritized spanners with lightness arbitrarily close to 1.

**Prioritized vs. Scaling Distortion** As mentioned above, one of the ingredients of our work is a *general reduction* relating the notions of prioritized distortion and scaling distortion. In fact, we show that prioritized distortion is essentially equivalent to a strong version of scaling distortion called *coarse* scaling distortion, in which for every point, the $1-\epsilon$ fraction of the farthest points from it are preserved with the desired distortion. We prove that any embedding with a given prioritized distortion $\alpha$ has coarse scaling distortion bounded by $O(\alpha(8/\epsilon))$. This result could be of independent interest; in particular, it shows that the results of [16] on distance oracles and embeddings have their scaling distortion counterparts (some of which were not known before). We further show a reduction in the opposite direction, informally, that given an embedding with coarse scaling distortion $\gamma$ there exists an embedding with prioritized distortion $\gamma(\mu(j))$, where $\mu$ is a function such that $\sum_i \mu(i) = 1$ (e.g. $\mu(j) = \frac{6}{(\pi\cdot j)^2}$). This result implies that all existing *coarse* scaling distortion results have priority distortion counterparts, thus improving few of the results of [16]. In particular, by applying a theorem of [3] we obtain prioritized embedding of arbitrary metric spaces into $l_p$ in dimension $O(\log n)$ and prioritized distortion $O(\log j)$, which is best possible.

**Outline and Techniques.** Our proof has the following high level approach; Given a graph and a ranking of its vertices, we first find a low weight spanner with prioritized distortion $\tilde{O}(\log j)/\rho$. We then apply the general reduction from prioritized distortion to scaling distortion to find a spanner with scaling distortion $\tilde{O}(\log(1/\epsilon))/\rho$. Finally, we use the result of [4] to find a spanning tree of this spanner with scaling distortion $O(1/\sqrt{\epsilon})$. We then conclude that the scaling distortion of the concatenated embeddings is roughly their product, which implies our main result of a spanning tree with lightness $1+\rho$ and scaling distortion $\tilde{O}(1/\sqrt{\epsilon})/\rho$.

Similarly, we can apply the probabilistic embedding of [4] to get a light counterpart, devising a distribution over spanning trees, each with lightness $1+\rho$, with (expected) scaling distotion $\mathrm{polylog}(1/\epsilon)/\rho$.

The main technical part of the paper is finding a light prioritized spanner. In a recent result [14] (following [18, 13]), it was shown that any graph on $n$ vertices admits a spanner with (worst-case) distortion $O(\log n)$ and with constant lightness. However, these constructions have no bound on the more refined notions of distortion. To obtain a prioritized distortion, we use a technique similar in spirit to [16]: group the vertices into $\log\log n$ sets according to their priority, the set $K_i$ will contain vertices with priority up to $2^{2^i}$. We then build a low weight spanner for each of these sets. As prioritized distortion guarantees a bound for *every pair* containing a high ranking vertex, we must augment the spanner of $K_i$ with shortest paths to all other vertices. Such a shortest path tree may have large weight, so we use an idea from [12] and apply an SLT rooted at $K_i$, which balances between the weight and the distortion from $K_i$.

The main issue with the construction described above is that the weight of the spanner in each phase can be proportional to that of the MST, but we have $\log\log n$ of those. Obtaining constant lightness, completely independent of $n$, requires a subtler argument. We use the fact that the weight of the light spanners in each phase come "mostly" from the MST, and then some additional weight. We ensure that all the spanners will have *the same* MST. Then we select the parameters carefully, so that the additional weights will be small enough to form converging sequences, without affecting the distortion by too much.

### 1.1 Related Work

Partial and scaling embeddings[4] have been studied in several papers [22, 1, 3, 12, 4, 5]. Some of the notable results are embedding arbitrary metrics into a distribution over trees [1] or into Euclidean space [3] with tight $O(\log(1/\epsilon))$ scaling distortion. These results imply constant average distortion and $O(q)$ bound on the $\ell_q$-distortion. In [4], an embedding into a single spanning tree with tight $O(1/\sqrt{\epsilon})$ scaling distortion is shown, which implies constant average distortion, but there is no guarantee on the weight of the tree.

Prioritized distortion embeddings were studied in [16], for instance they give an embedding of arbitrary metrics into a distribution over trees with prioritized distortion $O(\log j)$ and into Euclidean space with prioritized distortion $\tilde{O}(\log j)$.

Probabilistic embedding into trees [9, 10, 11, 19] and spanning trees [7, 15, 2, 6] has been intensively studied, and found numerous applications to approximation and online algorithms, and to fast linear system solvers. While our distortion guarantee does not match the best known worst-case bounds, which are $O(\log n)$ for arbitrary trees and $\tilde{O}(\log n)$ for spanning trees, we give the first probabilisitc embeddings into spanning trees with polylogarithmic scaling distortion in which all the spanning trees in the support of the distribution are light.

The paper [12] considers partial and scaling embedding into spanners, and show a general transformation from worst-case distortion to partial and scaling distortion. In particular, they show a spanner with $O(n)$ edges and $O(\log(1/\epsilon))$ scaling distortion. For a fixed $\epsilon > 0$, they also obtain a spanner with $O(n)$ edges, $O(\log(1/\epsilon))$ partial distortion and lightness $O(\log(1/\epsilon))$.[5] Note that these results fall short of achieving both constant average distortion and constant lightness.

## 2 Preliminaries

All the graphs $G = (V, E, w)$ we consider are undirected and weighted with nonnegative weights. We shall assume w.l.o.g that all edge weights are different. If it is not the case, then one can break ties in an arbitrary (but consistent) way. Note that under this assumption, the MST $T$ of $G$ is unique. The weight of a graph $G$ is $w(G) = \sum_{e\in E} w(e)$. Let $d_G$ be the shortest path metric on $G$. For a subset $K \subseteq V$ and $v \in V$ let $d_G(v, K) = \min_{u\in K}\{d_G(u, v)\}$. For $r \geq 0$ let $B_G(v, r) = \{u \in V : d_G(u, v) \leq r\}$ (we often omit the subscript when clear from context).

For a graph $G = (V, E)$ on $n$ vertices, a subgraph $H = (V, E')$ where $E' \subseteq E$ (with the induced weights) is called a *spanner* of $G$. We say that a pair $u, v \in V$ has *distortion* at most $t$ if

$$d_H(v, u) \leq t \cdot d_G(v, u) \ ,$$

(note that always $d_G(v, u) \leq d_H(v, u)$). If every pair $u, v \in V$ has distortion at most $t$, we say that the spanner $H$ has distortion $t$. Let $T$ be the (unique) MST of $G$, the *lightness* of $H$ is the ratio

[4]A partial embedding (introduced by [22] under the name *embedding with slack*) requires that for a *fixed* $0 < \epsilon < 1$, the distortion of all but an $\epsilon$-fraction of the pairs is bounded by the appropriate function of $\epsilon$.

[5]The original paper claims lightness $O(\log^2(1/\epsilon))$, but their proof in fact gives the improved bound.

between the weight of $H$ and the weight of the MST, that is $\Psi(H)=\frac{w(H)}{w(T)}$. We sometimes abuse notation and identify a spanner or a spanning tree with its set of edges.

**Prioritized Distortion.** Let $\pi=v_1,\ldots,v_n$ be a priority ranking (an ordering) of the vertices of $V$, and let $\alpha:\mathbb{N}\to\mathbb{R}_+$ be some monotone non-decreasing function. We say that $H$ has *prioritized distortion* $\alpha$ (w.r.t $\pi$), if for all $1\le j<i\le n$, the pair $v_j,v_i$ has distortion at most $\alpha(j)$.

**Scaling Distortion.** For $v\in V$ and $\epsilon\in(0,1)$ let $R(v,\epsilon)=\min\{r:|B(v,r)|\ge\epsilon n\}$. A vertex $u$ is called *$\epsilon$-far* from $v$ if $d(u,v)\ge R(v,\epsilon)$. Given a function $\gamma:(0,1)\to\mathbb{R}_+$, we say that $H$ has *scaling distortion* $\gamma$, if for every $\epsilon\in(0,1)$, there are at least $(1-\epsilon)\binom{|V|}{2}$ pairs that have distortion at most $\gamma(\epsilon)$. We say that $H$ has *coarse* scaling distortion $\gamma$, if every pair $v,u\in V$ such that both $u,v$ are $\epsilon/2$-far from each other, has distortion at most $\gamma(\epsilon)$.[6]

**Moments of Distortion.** For $1\le q\le\infty$, define the *$\ell_q$-distortion* of a spanner $H$ of $G$ as:

$$\mathrm{dist}_q(H,G)=\mathbb{E}\left[\left(\frac{d_H(u,v)}{d_G(u,v)}\right)^q\right]^{1/q},$$

where the expectation is taken according to the uniform distribution over $\binom{V}{2}$. The classic notion of *distortion* is expressed by the $\ell_\infty$-distortion and the *average distortion* is expressed by the $\ell_1$-distortion. The following was proved in [4].

**Lemma 1.** *([4]) Given a weighted graph $G=(V,E)$ on $n$ vertices, if a spanner $H$ has scaling distortion $\gamma$ then*

$$\mathrm{dist}_q(H,G)\le\left(2\int_{\frac{1}{2}\binom{n}{2}^{-1}}^{1}\gamma(x)^q dx\right)^{1/q}.$$

These notions of distortion apply for embedding of general metric spaces as well.

## 3 Light Spanner with Prioritized Distortion

In this section we prove that every graph admits a light spanner with bounded prioritized distortion.

**Theorem 1** (Prioritized Spanner)**.** *Given a graph $G=(V,E)$, a parameter $0<\rho<1$ and any priority ranking $v_1,v_2,\ldots,v_n$ of $V$, there exists a spanner $H$ with lightness $1+\rho$ and prioritized distortion $\tilde{O}(\log j)/\rho$.*

The main technical lemma is the following.

**Lemma 2.** *Given a graph $G=(V,E)$, a subset $K\subseteq V$ of size $k$, and a parameter $0<\delta<1$, there exists a spanner $H$ that* ***1)*** *contains the MST of $G$,* ***2)*** *has lightness $1+\delta$, and* ***3)*** *every pair in $K\times V$ has distortion $O(\log k/\delta)$.*

Before proving this lemma, we use it to prove Theorem 1.

[6] It can be verified that coarse scaling distortion $\gamma$ implies scaling distortion $\gamma$.

*Proof.* (of Theorem 1) For every $1 \le i \le \lceil \log\log n \rceil$ let $K_i = \left\{ v_j \ : \ j \le 2^{2^i} \right\}$. Let $H_i$ be the spanner given by Lemma 2 with respect to the set $K_i$ and the parameter $\delta_i = \rho/i^2$. Hence $H_i$ has $1+\rho/i^2$ lightness and $O\left(\frac{\log|K_i|}{\delta_i}\right) = O(2^i \cdot i/\rho)$ distortion for pairs in $K_i \times V$. Let $H = \bigcup_i H_i$ be the union of all these spanners (that is, the graph containing every edge of every one of these spanners). As each $H_i$ contains the unique MST of $G$, it holds that

$$\Psi(H) \ \le \ 1 + \sum_{i \ge 1} \rho/i^2 = 1 + O(\rho) \ .$$

To see the prioritized distortion, let $v_j, v_r \in V$ be such that $j < r$, and let $1 \le i \le \lceil \log\log n \rceil$ be the minimal index such that $v_j \in K_i$. Note that $2^{2^{i-1}} \le j$, and in particular $2^{i-1} \le \log j$ (with the exception of $j = 1$, which we may ignore). This implies that

$$\begin{aligned} d_H(v_j, v_r) \ &\le \ d_{H_i}(v_j, v_r) \ \le \ O(2^i \cdot i^2/\rho) \cdot d_G(v_j, v_r) \\ &\le \ \tilde{O}(\log j)/\rho \cdot d_G(v_j, v_r) \ . \end{aligned}$$

as required. □

### 3.1 Proof of Lemma 2

The construction of the spanner that fullfil the properties promised in Lemma 2 is as follows. First we use the spanner of [14] to get a spanner with lightness $O(1)$ and distortion $O(\log k)$ over pairs in $K \times K$. Then, by combining this spanner with the SLT by [21] we expand the $O(\log k)$ distortion guarantee to all pairs in $K \times V$, while the lightness is still $O(1)$. Finally, we use general reduction (Theorem 2), that reduces the weight of a spanner while increasing its distortion. By applying the reduction, we get a spanner with $1+\rho$ lightness while paying additional factor of $\rho$ in the distortion.

We begin by describing the general reduction.

**Theorem 2.** *Let $G = (V, E)$ be a graph, $0 < \delta < 1$ a parameter and $t : \binom{V}{2} \to \mathbb{R}_+$ some a function. Suppose that for every weight function $w : E \to \mathbb{R}_+$ there exist a spanner $H$ with lightness $\ell$ such that every pair $u, v$ suffers distortion at most $t(u, v)$. Then for every weight function $w$ there exist a spanner $H$ with lightness $1 + \delta\ell$ and such that every pair $u, v$ suffers distortion at most $t(u, v)/\delta$. Moreover, $H$ contains the MST of $G$ with respect to $w$.*

*Proof.* Let $T$ be the MST of $G$ with respect to $w$. Set $w' : E \to \mathbb{R}_+$ to be a new weight function $w'(e) = \begin{cases} \delta \cdot w(e) & e \in T \\ w(e) & e \notin T \end{cases}$, that is, we multiply the weight of all the MST edges by $\delta$. Let $G'$ be the graph $G$ with the weight function $w'$. Note that $T$ is also the MST of $G'$ (as every $e \notin T$ is still the heaviest edge in some cycle). Using our assumption, let $H' = (V, E_{H'}, w')$ be a spanner of $G'$ with stretch $t$ and lightness $\ell$. Set $H = (V, E_{H'} \cup T, w)$ as a spanner of $G$. The edge set of $H$ consist of the edges of $H'$ union the MST edges with the original weight function $w$.

As the weight of the non-MST edges did not changed, we have

$$\begin{aligned} w(E_H) \ &= \ w(T) + w'(E_{H'} \setminus T) \le w(T) + w'(E_{H'}) \\ &\le \ w(T) + \ell \cdot w'(T) = (1 + \delta\ell) \cdot w(T) \ , \end{aligned}$$

in particular the lightness of $H$ is at most $1 + \delta\ell$.

To bound the distortion, consider an arbitrary pair of vertices $u, v \in V$. Let $P_{u,v}$ be the shortest path from $u$ to $v$ in $H'$. As for each edge $e \in P_{u,v}$, $w(e) \le w'(e)/\delta$ , we have

$$d_H(u,v) \le \sum_{e \in P_{u,v}} w(e) \le \frac{1}{\delta} \cdot \sum_{e \in P_{u,v}} w'(e) = \frac{1}{\delta} \cdot d_{H'}(u,v) \le \frac{t(u,v)}{\delta} d_{G'}(u,v) \le \frac{t(u,v)}{\delta} d_G(u,v) \ ,$$

as required. □

In a recent work, Chechik and Wulff-Nilsen achieved the following result:

**Theorem 3** ([14]). *For every weighted graph $G = (V, E, w)$ and parameters $k \ge 1$ and $0 < \epsilon < 1$, there exist a polynomial time algorithm that constructs a spanner with distortion $(2t-1)(1+\epsilon)$ and lightness $n^{1/t} \cdot \mathrm{poly}(\frac{1}{\epsilon})$.*

Note that for an $n$-vertex graph with parameters $t = \log n, \epsilon = 1$, they get a spanner with distortion $O(\log n)$ and constant lightness. However, their construction does not provide lightness arbitrarily close to 1.

A tree $\mathcal{T} = (V', E', w')$ is called a *Steiner tree* for a graph $G = (V, E, w)$ if (1) $V \subseteq V'$, and (2) for any pair of vertices $u, v \in V$ it holds that $d_{\mathcal{T}}(u,v) \ge d_G(u,v)$. The *minimum Steiner tree* $T$ of $G$, denoted $SMT(G)$, is a Steiner tree of $G$ with minimum weight. It is well-known that for any graph $G$, $w(SMT(G)) \ge \frac{1}{2} MST(G)$. (See, e.g., [20], Section 10.)

We will use [14] spanner to construct a spanner with $O(1)$ lightness and distortion $O(\log k)$ over pairs in $K \times K$. Let $G_k = (K, \binom{K}{2}, d_G)$ be the complete graph over the terminal set $K$ with weights according to the shortest path metric in $G$. Let $T_k$ be the MST of $G_k$. Note that the MST $T$ of $G$ is a Steiner tree of $G_k$, hence $w_k(T_k) \le 2 \cdot w_k(SMT(G_k)) \le 2 \cdot w(T)$ .

Using Theorem 3, let $H_k$ be a spanner of $G_k$ with weight $O(w(T))$ (constant lightness) and distortion $O(\log k)$. For a pair of vertices $u, v \in V$, let $P_{uv}$ denote a shortest path between $u$ and $v$. Let $H' = (V, E', w)$ be a subgraph of $G$ with the set of edges $E' = \cup_{\{u,v\} \in E_k} P_{uv}$ (i.e. for every edge $\{u, v\}$ in $H_k$, we take some shortest path from $u$ to $v$). It holds that,

$$w(H') \le \sum_{\{u,v\} \in E_k} w(P_{uv}) = \sum_{e \in E_k} w_k(e) = O(1) \cdot w_k(T_k) = O(1) \cdot w(T) \ .$$

Moreover, for every pair $u, v \in K$,

$$d_{H'}(u,v) \le d_H(u,v) \le O(\log k) \cdot d_G(u,v) \ . \tag{1}$$

Now we extend $H'$ so that every pair in $K \times V$ will suffer distortion at most $O(\log k)$. To this end, we use the following lemma regarding shallow light trees (SLT) is implicitly proved in [21, 8].

**Lemma 3.** *Given a graph $G = (V, E)$, a parameter $\alpha > 1$, and a subset $K \subseteq V$, there exists a spanner $S$ of $G$ with lightness $1 + \frac{2}{\alpha - 1}$, and for any vertex $u \in V$, $d_S(u, K) \le \alpha \cdot d_G(u, K)$.*

Let $S$ be the spanner of Lemma 3 with respect to the set $K$ and parameter $\alpha = 2$. Define $H$" as the union of $H'$ and $S$. As both $H'$ and $S$ have constant lightness, so does $H$". It remains to bound the distortion of an arbitrary pair $v \in K$ and $u \in V$. Let $k_u \in K$ be the closest vertex to $u$ among the vertices in $K$ with respect to the distances in the spanner $S$. By the assertion of Lemma 3,

$$d_S(u, k_u) = d_S(u, K) \le 2 \cdot d_G(u, K) \le 2 \cdot d_G(u, v) \ . \tag{2}$$

Using the triangle inequality,

$$d_G(v,k_u) \leq d_G(v,u) + d_G(u,k_u) \leq d_G(v,u) + d_S(u,k_u) \overset{(2)}{\leq} 3 \cdot d_G(v,u) \ . \tag{3}$$

Since both $v, k_u \in K$ it follows that

$$d_{H'}(v,k_u) \overset{(1)}{\leq} O(\log k) \cdot d_G(v,k_u) \overset{(3)}{\leq} O(\log k) \cdot d_G(v,u) \ . \tag{4}$$

We conclude that

$$d_{H"}(v,u) \leq d_{H'}(v,k_u) + d_S(k_u,u) \overset{(2)\wedge(4)}{\leq} O(\log k) \cdot d_G(v,u) \ .$$

We showed a polynomial time algorithm that given a weighted graph $G = (V,E,w)$ and a subset $K \subseteq V$ of size $k$ constructs a spanner $H$ with lightness $O(1)$ and such that every pair in $K \times V$ has distortion at most $O(\log k)$. Now Theorem 2 implies Lemma 2.

# 4 Prioritized Distortion vs. Coarse Scaling Distortion

In this section we study the relationship between the notions of prioritized and scaling distortion. We show that there is a reduction that allows to transform embeddings with prioritized distortion into embeddings with coarse scaling distortion, and vice versa. We start with the direction that is used for our main result, showing that prioritized distortion implies scaling distortion.

For two metric spaces $(X, d_X)$, $(Y, d_Y)$ and a non-contractive embedding $f : X \to Y$,[7] the distortion of a pair $x, y \in X$ under $f$ is defined as $\frac{d_Y(f(x),f(y))}{d_X(x,y)}$.

**Theorem 4.** *Let $(X, d_X)$, $(Y, d_Y)$ be metric spaces, then there exists a priority ranking $x_1, \ldots, x_n$ of the points of $X$ such that the following holds: If there exists a non-contractive embedding $f : X \to Y$ with (monotone non-decreasing) prioritized distortion $\alpha$, then $f$ has coarse scaling distortion $O(\alpha(8/\epsilon))$.*

The basic idea of the proof is to choose the priorities so that for every $\epsilon$, every $v \in X$ has a representative $v'$ of sufficiently high priority within distance $\approx R(v,\epsilon)$. Then for any $u \in X$ which is $\epsilon$-far from $v$, we can use the low distortion guarantee of $v'$ with both $v$ and $u$ via the triangle inequality. To this end, we employ the notion of a *density net* due to [12], who showed that a greedy construction provides such a net.

**Definition 1** (Density Net)**.** *Given a metric space $(X, d)$ and a parameter $0 < \epsilon < 1$, an* $\epsilon$-density-net *is a set $N \subseteq X$ such that:* ***1)*** *for all $v \in X$ there exists $u \in N$ with $d(v,u) \leq 2R(v,\epsilon)$ and* ***2)*** *$|N| \leq \frac{1}{\epsilon}$.*

*Proof.* (of Theorem 4) We begin by describing the desired priority ranking of $X$. For every integer $1 \leq i \leq \lceil \log n \rceil$ let $\epsilon_i = 2^{-i}$, and let $N_i \subseteq X$ be an $\epsilon_i$-density-net in $X$. Take any priority ranking of $X$ satisfying that every point $v \in N_i$ has priority at most $\left|\bigcup_{j=1}^{i} N_j\right| \leq \sum_{j=1}^{i} |N_j|$. As for any $j$, $|N_j| \leq \frac{1}{\epsilon_j} = 2^j$, each point in $N_i$ has priority at most $\sum_{j=1}^{i} \frac{1}{\epsilon_j} \leq \sum_{j=1}^{i} 2^j < 2^{i+1}$.

[7] An embedding $f$ is non-contractive if for every $x, y \in X$, $d_Y(f(x), f(y)) \geq d_X(x, y)$.

Let $f : X \to Y$ be some non-contractive embedding with prioritized distortion $\alpha$ with respect to the priorities we defined. Fix some $\epsilon \in (0,1)$ and a pair $v, u \in V$ so that $u$ is $\epsilon$-far from $v$. Let $i$ be the minimal integer such that $\epsilon_i \le \epsilon$ (note that we may assume $1 \le i \le \lceil \log n \rceil$, because there is nothing to prove for $\epsilon < 1/n$). By Definition 1 we can take $v' \in N_i$ such that $d(v, v') \le 2R(v, \epsilon_i)$. As $u$ is $\epsilon$-far from $v$, it holds that

$$d_X(v, v') \le 2R(v, \epsilon_i) \le 2R(v, \epsilon) \le 2d_X(v, u) \ . \tag{5}$$

In particular, by the triangle inequality,

$$d_X(u, v') \le d_X(u, v) + d_X(v, v') \overset{(5)}{\le} 3d_X(u, v) \ . \tag{6}$$

The priority of $v'$ is at most $2^{i+1}$, hence

$$\begin{aligned} & d_Y(f(v), f(u)) \\ & \quad \le \quad d_Y(f(v), f(v')) + d_Y(f(v'), f(u)) \\ & \quad \le \quad \alpha(2^{i+1}) \cdot d_X(v, v') + \alpha(2^{i+1}) \cdot d_X(v', u) \\ & \quad \overset{(5)\wedge(6)}{\le} \quad 5\alpha(2/\epsilon_i) \cdot d_X(v, u) \ . \end{aligned}$$

By the minimality of $i$ it follows that $1/\epsilon_i \le 2/\epsilon$, and since $\alpha$ is monotone

$$d_Y(f(v), f(u)) \le 5\alpha(2/\epsilon_i) \cdot d_X(v, u) \le 5\alpha(4/\epsilon) \cdot d_X(v, u) \ ,$$

as required. Since we desire distortion guarantee for pairs that are $\epsilon/2$-far, the distortion becomes $O(\alpha(8/\epsilon))$.

□

Combining Theorem 1 and Theorem 4 we obtain the following.

**Theorem 5.** *For any parameter $0 < \rho < 1$, any graph contains a spanner with coarse scaling distortion $\tilde{O}(\log(1/\epsilon))/\rho$ and lightness $1 + \rho$.*

**Remark 1.** *By Lemma 1 it follows that this spanner has $\ell_q$-distortion $\tilde{O}(q)/\rho$ for any $1 \le q < \infty$.*

*We can also obtain a spanner with both scaling distortion and prioritized distortion simultaneously, where the priority is with respect to an arbitrary ranking $\pi = v_1, \ldots, v_n$. To achieve this, one may define a ranking which interleaves $\pi$ with the ranking generated in the proof of Theorem 4. We leave the details to the reader.*

We now turn to show that coarse scaling distortion implies prioritized distortion.

**Theorem 6.** *Let $\mu : N \to \mathbb{R}^+$ be a non-increasing function such that $\sum_{i \ge 1} \mu(i) = 1$. Let $\mathcal{Y}$ be a family of finite metric spaces, and assume that for every finite metric space $(Z, d_Z)$ there exists a non-contractive embedding $f_Z : Z \to Y_Z$, where $(Y_Z, d_{Y_Z}) \in \mathcal{Y}$, with (monotone non-increasing) coarse scaling distortion $\gamma$. Then, given a finite metric space $(X, d_X)$ and a priority ranking $x_1, \ldots, x_n$ of the points of $X$, there exists an embedding $f : X \to Y$, for some $(Y, d_Y) \in \mathcal{Y}$, with (monotone non-decreasing) prioritized distortion $\gamma(\mu(i))$.*

*Proof.* Given the metric space $(X, d_X)$ and a priority ranking $x_1, \ldots, x_n$ of the points of $X$, let $\delta = \min_{i \neq j} d_X(x_i, x_j)/2$. We define a new metric space $(Z, d_Z)$ as follows. For every $1 \leq i \leq n$, every point $x_i$ is replaced by a set $X_i$ of $|X_i| = \lceil \mu(i)n \rceil$ points, and let $Z = \bigcup_{i=1}^n X_i$. For every $u \in X_i$ and $v \in X_j$ define $d_Z(u,v) = d_X(x_i, x_j)$ when $i \neq j$, and $d_Z(u,v) = \delta$ otherwise. Observe that $|Z| = \sum_{i=1}^n |X_i| \leq \sum_{i=1}^n (\mu(i)n + 1) \leq 2n$.

We now use the embedding $f_Z : Z \to Y_Z$ with coarse scaling distortion $\gamma$, to define an embedding $f : X \to Y_Z$, by letting for every $1 \leq i \leq n$, $f(x_i) = f_Z(u_i)$ for some (arbitrary) point $u_i \in X_i$. By construction of $Z$, for every $j > i$, we have that $X_i \subseteq B(u_i, d_Z(u_i, u_j)) \cap B(u_j, d_Z(u_i, u_j))$. As $|X_i| \geq \mu(i)n \geq \frac{\mu(i)}{2}|Z|$, it holds that $u_i, u_j$ are $\epsilon/2$-far from each other for $\epsilon = \mu(i)$. This implies that $\frac{d_{Y_Z}(f(x_i), f(x_j))}{d_X(x_i, x_j)} = \frac{d_{Y_Z}(f_Z(u_i), f_Z(u_j))}{d_Z(u_i, u_j)} \leq \gamma(\mu(i))$. □

It follows from a result of [16] that the convergence condition on $\mu$ in the above theorem is necessary. We note that this reduction can also be applied to cases where the coarse scaling embedding is only known for a class of metric spaces (rather than all metrics), as long as the transformation needed for the proof can be made so that the resulting new space is still in the class. This holds for most natural classes. We leave the details for the full version of the paper.

The reduction implies that all existing *coarse* scaling distortion results have priority distortion counterparts, thus improving few of the results of [16][8]. In particular, by applying a theorem of [3] we get the following:

**Theorem 7.** *For every $1 \leq p \leq \infty$ and every finite metric space $(X, d_X)$, and a priority ranking of $X$, there exists an embedding with prioritized distortion $O(\log j)$ into $l_p^{O(\log |X|)}$.*

**Remark 2.** *The proof of Theorem 4 provides an even stronger conclusion, that any pair $u, v \in X$ such that one is $\epsilon/2$-far from the other, has the claimed distortion bound. While in the original definition of coarse scaling both points are required to be $\epsilon/2$-far from each other, it is often the case that we acheive the stronger property. Yet, in some of the cases in previous work the weaker definition seemed to be of importance. Combining Theorem 4 and Theorem 6, we infer that essentially any coarse scaling embedding can have such a one-sided guarantee, with a slightly worse dependence on $\epsilon$, as claimed in the following corollary.*

**Corollary 4.** *Fix a metric space $(X, d)$ on $n$ points. Let $\mathcal{Y}$ be a family of finite metric spaces as in Theorem 6. Then there exists an embedding $f : X \to Y$, for some $(Y, d_Y) \in \mathcal{Y}$, with (monotone non-decreasing)* one-sided *coarse scaling distortion $O(\gamma(\mu(8/\epsilon)))$, where $\mu : N \to \mathbb{R}^+$ is a non-increasing function such that $\sum_{i \geq 1} \mu(i) = 1$. .*

*Proof.* By the condition of Theorem 6, there exists $(Y, d_Y) \in \mathcal{Y}$ so that $X$ embeds to $Y$ with coarse scaling distortion $\gamma(\epsilon)$. According to Theorem 6, there is an embedding $f$ with prioritized distortion $\gamma(\mu(i)))$ (w.r.t to any fixed priority ranking $\pi$). We pick $\pi$ to be the ordering required by Theorem 4, and conclude that $f$ has strong coarse scaling distortion $O(\gamma(\mu(8/\epsilon)))$. □

# 5 A Light Tree with Constant Average Distortion

Here we prove our main theorem on finding a light spanning tree with constant average distortion. Later on we show a probabilistic embedding into a distribution of light spanning trees with improved

[8]It is also worth noting that the reduction also implies that coarse partial embedding results can be translated into bounds on terminal distortion [17].

bound on higher moments of the distortion.

**Theorem 8.** *For any parameter $0 < \rho < 1$, any graph contains a spanning tree with scaling distortion $\tilde{O}(\sqrt{1/\epsilon})/\rho$ and lightness $1+\rho$.*

It follows from Lemma 1 that the average distortion of the spanning tree obtained is $O(1/\rho)$. Moreover, the $\ell_q$-distortion is $O(1/\rho)$ for any fixed $1 \le q < 2$, $\tilde{O}\left(\log^{1.5} n\right)/\rho$ for $q = 2$, and $\tilde{O}(n^{1-2/q})/\rho$ for any fixed $2 < q < \infty$.

We will need the following simple lemma, that asserts the scaling distortion of a composition of two maps is essentially the product of the scaling distortions of these maps.[9]

**Lemma 5.** *Let $(X, d_X)$, $(Y, d_Y)$ and $(Z, d_Z)$ be metric spaces. Let $f : X \to Y$ (respectively, $g : Y \to Z$) be a non-contractive onto embedding with scaling distortion $\alpha$ (resp., $\beta$). Then $g \circ f$ has scaling distortion $\alpha(\epsilon/2) \cdot \beta(\epsilon/2)$.*

*Proof.* Let $n = |X|$. Let $\text{dist}_f(v,u) = \frac{d_Y(f(v),f(u))}{d_X(v,u)}$ be the distortion of the pair $u, v \in X$ under $f$, and similarly let $\text{dist}_g(v,u) = \frac{d_Z(g(f(v)),g(f(u)))}{d_Y(f(v),f(u))}$. Fix some $\epsilon \in (0,1)$. We would like to show that at most $\epsilon \cdot \binom{n}{2}$ pairs suffer distortion greater than $\alpha(\epsilon/2)\cdot\beta(\epsilon/2)$ by $g \circ f$. Let $A = \left\{\{v,u\} \in \binom{X}{2} : \text{dist}_f(v,u) > \alpha(\epsilon/2)\right\}$ and $B = \left\{\{v,u\} \in \binom{X}{2} : \text{dist}_g(v,u) > \beta(\epsilon/2)\right\}$. By the bound on the scaling distortions of $f$ and $g$, it holds that $|A \cup B| \le |A| + |B| \le \epsilon \cdot \binom{n}{2}$. Note that if $\{v,u\} \notin A \cup B$ then

$$\begin{aligned} \frac{d_Z\left(g(f(v)), g(f(u))\right)}{d_X\left(v,u\right)} &= \text{dist}_f(v,u)\cdot \text{dist}_g(v,u) \\ &\le \alpha(\epsilon/2)\cdot\beta(\epsilon/2)\ , \end{aligned}$$

which concludes the proof. □

We will also need the following result, that was proved in [4].

**Theorem 9** ([4])**.** *Any graph contains a spanning tree with scaling distortion $O(\sqrt{1/\epsilon})$.*

Now we can prove the main result.

*Proof.* (of Theorem 8) Let $H$ be the spanner given by Theorem 5. Let $T$ be a spanning tree of $H$ constructed according to Theorem 9. By Lemma 5, $T$ has scaling distortion $O(\sqrt{1/\epsilon}) \cdot \tilde{O}(\log(1/\epsilon))/\rho = \tilde{O}(\sqrt{1/\epsilon})/\rho$ with respect to the distances in $G$. The lightness follows as $\Psi(T) \le \Psi(H) \le 1+\rho$.

□

**Random Tree Embedding.** We also derive a result on probabilistic embedding into light spanning trees with scaling distortion. That is, the embedding construct a distribution over spanning tree so that each tree in the support of the distribution is light. In such probabilistic embeddings [9] into a family $\mathcal{Y}$, each embedding $f = f_Y : X \to Y$ (for some $(Y, d_Y) \in \mathcal{Y}$) in the support of the distribution is non-contractive, and the distortion of the pair $u, v \in X$ is defined as $\mathbb{E}_Y\left[\frac{d_Y(f(u),f(v))}{d_X(u,v)}\right]$.

[9] Note that this is not true for the average distortion – one may compose two maps with constant average distortion and obtain a map with $\Omega(n)$ average distortion.

The prioritized and scaling distortions are defined accordingly. We make use of the following result from [4].[10]

**Theorem 10.** *([4]) Every weighted graph $G$ embeds into a distribution over spanning trees with coarse scaling distortion $\tilde{O}(\log^2(1/\epsilon))$.*

We note that the distortion bound on the composition of maps in Lemma 5 also holds whenever $g$ is a random embedding, and we measure the scaling expected distortion. Thus, following the same lines as in the proof of Theorem 8, (while using Theorem 10 instead of Theorem 9), we obtain the following.

**Theorem 11.** *For any parameter $0 < \rho < 1$ and any weighted graph $G$, there is an embedding of $G$ into a distribution over spanning trees with scaling distortion $\tilde{O}(\log^3(1/\epsilon))/\rho$, such that every tree $T$ in the support has lightness $1 + \rho$.*

It follows from Lemma 1 that the $\ell_q$-distortion is $O(1/\rho)$, for every fixed $q \geq 1$.

# 6 Lower Bound on the Trade-off between Lightness and Average Distortion

In this section, we give an example of a graph for which any spanner with lightness $1+\rho$ has average distortion $\Omega(1/\rho)$ (of course this bound holds for the $\ell_q$-distortion as well). This shows that our results are tight [11].

**Lemma 6.** *For any $n \geq 32$ and $\rho \in [1/n, 1/32]$, there is a graph $G$ on $n+1$ vertices such that any spanner $H$ of $G$ with lightness at most $1+\rho$ has average distortion at least $\Omega(1/\rho)$.*

*Proof.* We define the graph $G = (V, E)$ as follows. Denote $V = \{v_0, v_1, \ldots, v_n\}$, $E = \binom{V}{2}$, and the weight function $w$ is defined as follows.

$$w(\{v_i, v_j\}) = \begin{cases} 1 & \text{if } |i-j| = 1 \\ 2 & \text{otherwise .} \end{cases}$$

I.e., $G$ is a complete graph of size $n+1$, where the edges $\{v_i, v_{i+1}\}$ have unit weight and induce a path of length $n$, and all non-path edges have weight 2. Clearly, the path is the MST of $G$ of weight $n$. Let $k = \lceil \rho n \rceil$. Let $H$ be some spanner of $G$ with lightness at most $1 + \rho \leq \frac{n+k}{n}$, in particular, $w(H) \leq n + k$. Clearly $H$ has at least $n$ edges (to be connected). Let $q$ be the number of edges of weight 2 contained in $H$. Then $w(H) \geq (n-q)\cdot 1 + q \cdot 2 = n + q$. Therefore $q \leq k$.

Let $S$ be the set of vertices which are incident on an edge of weight 2 in $H$. Then $|S| \leq 2q \leq 2k$. Let $\delta = \frac{1}{32\rho}$. For any $v \in S$, let $N_v \subseteq V$ be the set of vertices that are connected to $v$ via a path of length at most $\delta$ in $H$, such that this path consists of weight 1 edges only. Necessarily, for any $v \in S$, $|N_v| \leq 2\delta + 1$. Let $N = \bigcup_{v\in S} N_v$, it holds that $|N| \leq 2k \cdot (2\delta + 1) \leq 4\rho n(\frac{1}{16\rho} + 1) \leq \frac{n}{4} + \frac{n}{8} = \frac{3n}{8}$. Let $\bar{N} = V \setminus N$.

[10] The fact the embedding yields coarse scaling distortion is implicit in their proof.

[11] We also mention that in general the average distortion of a spanner cannot be arbitrarily close to 1, unless the spanner is extremely dense. E.g., when $G$ is a complete graph, any spanner with lightness at most $n/2$ will have average distortion at least $3/2$.

Consider $u \in \bar{N}$. By definition of $N$ every weight 2 edge is further than $\delta$ steps away from $u$ in $H$. It follows that there are at most $2\delta + 1$ vertices within distance at most $\delta$ from $u$ (in $H$). Let $F_u = \{v \in V \ : \ d_H(u,v) > \delta\}$. It follows that $|F_u| \geq n - 2\delta - 1$. Note that for any $v \in F_u$, the distortion of the pair $\{u,v\}$ is at least $\frac{\delta}{2}$. Hence, we obtain that

$$\begin{aligned} \sum_{\{v,u\} \in \binom{V}{2}} \frac{d_H(v,u)}{d_G(v,u)} &\geq \frac{1}{2} \sum_{u \in \bar{N}} \sum_{v \in F_u} \frac{d_H(v,u)}{d_G(v,u)} \\ &\geq \frac{5n}{16} \cdot (n - 2\delta - 1) \cdot \frac{\delta}{2} \\ &\geq \frac{5n}{16} \cdot \frac{7n}{8} \cdot \frac{1}{64\rho} \ . \end{aligned}$$

Finally,

$$\begin{aligned} \mathrm{dist}_1(H,G) &= \frac{1}{\binom{n+1}{2}} \sum_{\{v,u\} \in \binom{V}{2}} \frac{d_H(v,u)}{d_G(v,u)} \\ &\geq \frac{n}{n+1} \cdot \frac{35}{64} \cdot \frac{1}{64\rho} \\ &\geq \frac{1}{128\rho} \ . \end{aligned}$$

□


# 7 Acknowledgements

We are grateful to Michael Elkin for fruitful discussions.


# References


[1] Ittai Abraham, Yair Bartal, Hubert T.-H. Chan, Kedar Dhamdhere, Anupam Gupta, Jon M. Kleinberg, Ofer Neiman, and Aleksandrs Slivkins. Metric embeddings with relaxed guarantees. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005), 23-25 October 2005, Pittsburgh, PA, USA, Proceedings*, pages 83–100, 2005.

[2] Ittai Abraham, Yair Bartal, and Ofer Neiman. Nearly tight low stretch spanning trees. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '08, pages 781–790, Washington, DC, USA, 2008. IEEE Computer Society.

[3] Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. *Advances in Mathematics*, 228(6):3026 – 3126, 2011.

[4] Ittai Abraham, Yair Bartal, and Ofer Neiman. Embedding metrics into ultrametrics and graphs into spanning trees with constant average distortion. *SIAM J. Comput.*, 44(1):160–192, 2015.

[5] Ittai Abraham, Yair Bartal, Ofer Neiman, and Leonard J. Schulman. Volume in general metric spaces. *Discrete & Computational Geometry*, 52(2):366–389, 2014.

[6] Ittai Abraham and Ofer Neiman. Using petal-decompositions to build a low stretch spanning tree. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing*, STOC '12, pages 395–406, New York, NY, USA, 2012. ACM.

[7] Noga Alon, Richard M. Karp, David Peleg, and Douglas West. A graph-theoretic game and its application to the $k$-server problem. *SIAM J. Comput.*, 24(1):78–100, 1995.

[8] B. Awerbuch, A. Baratz, and D. Peleg. Efficient broadcast and light-weight spanners. Technical Report CS92-22, The Weizmann Institute of Science, Rehovot, Israel., 1992.

[9] Yair Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, pages 184–, Washington, DC, USA, 1996. IEEE Computer Society.

[10] Yair Bartal. On approximating arbitrary metrices by tree metrics. In *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, STOC '98, pages 161–168, New York, NY, USA, 1998. ACM.

[11] Yair Bartal. Graph decomposition lemmas and their role in metric embedding methods. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 89–97, 2004.

[12] T.-H. Hubert Chan, Michael Dinitz, and Anupam Gupta. Spanners with slack. In *Proceedings of the 14th Conference on Annual European Symposium - Volume 14*, ESA'06, pages 196–207, London, UK, UK, 2006. Springer-Verlag.

[13] Barun Chandra, Gautam Das, Giri Narasimhan, and José Soares. New sparseness results on graph spanners. In *Proceedings of the Eighth Annual Symposium on Computational Geometry*, SCG '92, pages 192–201, New York, NY, USA, 1992. ACM.

[14] Shiri Chechik and Christian Wulff-Nilsen. Near-optimal light spanners. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2016, Arlington, VA, USA, January 10-12, 2016*, pages 883–892, 2016.

[15] Michael Elkin, Yuval Emek, Daniel A. Spielman, and Shang-Hua Teng. Lower-stretch spanning trees. *SIAM Journal on Computing*, 38(2):608–628, 2008.

[16] Michael Elkin, Arnold Filtser, and Ofer Neiman. Prioritized metric structures and embedding. In *Proceedings of the 47th ACM Symposium on Theory of Computing*, STOC, 2015.

[17] Michael Elkin, Arnold Filtser, and Ofer Neiman. Terminal embeddings. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2015, August 24-26, 2015, Princeton, NJ, USA*, pages 242–264, 2015.

[18] Michael Elkin, Ofer Neiman, and Shay Solomon. Light spanners. In *Automata, Languages, and Programming - 41st International Colloquium, ICALP 2014, Copenhagen, Denmark, July 8-11, 2014, Proceedings, Part I*, pages 442–452, 2014.

[19] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, STOC '03, pages 448–455, New York, NY, USA, 2003. ACM.

[20] E. N. Gilbert and H. O. Pollak. Steiner minimal trees. *SIAM Journal on Applied Mathematics*, 16(1):1–29, Jan 1968.

[21] Samir Khuller, Balaji Raghavachari, and Neal E. Young. Balancing minimum spanning and shortest path trees. In *SODA*, pages 243–250, 1993.

[22] Jon Kleinberg, Aleksandrs Slivkins, and Tom Wexler. Triangulation and embedding using small sets of beacons. *Journal of the ACM*, 56(6):32:1–32:37, September 2009.