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

Yair Bartal * Arnold Filtser† Ofer Neiman‡

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* $O(1/\rho^2)$.[1]

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.

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) [20, 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).[2] 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 near optimal $(1+\rho)$ lightness that has $O(1/\rho^2)$ average distortion over all pairwise distances. One may wonder if it is indeed necessary to pay the slight extra cost in weight in order to preserve the average distortion. We show that it is indeed the case by exhibiting a lower bound on the tradeoff between lightness and average distortion, showing that the average distortion must

*School of Engineering and Computer Science, Hebrew University. Email: yair@cs.huji.ac.il. Supported in part by a grant from the Israeli Science Foundation (1609/11)

†Department of Computer Science, Ben-Gurion University of the Negev. Email: arnoldf@cs.bgu.ac.il. Partially supported by the Lynn and William Frankel Center for Computer Sciences, ISF grant No. (523/12) and by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°303809.

‡Department of Computer Science, Ben-Gurion University of the Negev. Email: neimano@cs.bgu.ac.il. Supported in part by ISF grant No. (523/12) and by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°303809.

[1] Recently we improved the average distortion to an optimal $O(1/\rho)$, see Section 7.

[2] Distortion is sometimes referred to as stretch.

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 [21, 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 [21][3], requires that for *every* $0 < \epsilon < 1$, the distortion of all but an ϵ-fraction of the pairs is bounded by the appropriate function of ϵ. 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[4] $\tilde{O}(1/\sqrt{\epsilon})/\rho^2$, which is nearly tight as a function of ϵ [4].

We also obtain a probabilistic embedding devising a distribution over (light) spanning trees with $\text{polylog}(1/\epsilon)/\rho^2$ 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 $\tilde{O}(\log^2 j)/\rho^2$. 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 $\tilde{O}(\log^2(1/\epsilon))/\rho^2$, and thus average distortion $O(1/\rho^2)$. 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 ℓ_q-distortion). Moreover, the worst-case distortion is only polylogarithmic in n. We note that all of our results admit deterministic polynomial time algorithms.

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 α 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 γ there exists an embedding with prioritized distortion $\gamma(\mu(j))$, where μ 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 $\tilde{O}(\log^2 j)/\rho^2$. We then apply the general reduction from prioritized distortion to scaling distortion to find a spanner with scaling distortion $\tilde{O}(\log^2(1/\epsilon))/\rho^2$. 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 $\tilde{O}(1/\sqrt{\epsilon})/\rho^2$.

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 $\text{polylog}(1/\epsilon)/\rho^2$.

The main technical part of the paper is finding a light prioritized spanner. In [13] it was shown that any graph on n vertices admits a spanner with (worst-

[3]Originally coined gracefully degrading embedding.

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

case) distortion $O(\log^2 n)$ and with constant lightness. This result was recently improved by [18] to distortion $O(\log^2 n/\log\log n)$. 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 each spanner and each SLT 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 and SLT's come "mostly" from the MST, and then some additional weight. First we change the constructions to ensure that each spanner and each SLT will have *the same* MST (a priori every set K_i may have a different MST spanning it, and each of the SLT's rooted at K_i may use different edges). 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[5] have been studied in several papers [21, 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 ℓ_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 $O(\log(1/\epsilon))$.[6] 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 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* α

[5] A partial embedding (introduced by [21] under the name *embedding with slack*) requires that for a *fixed* $0 < \epsilon < 1$, the distortion of all but an ϵ-fraction of the pairs is bounded by the appropriate function of ϵ.

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

(w.r.t π), 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 *ϵ-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* γ, 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 γ, if every pair $v, u \in V$ such that both u, v are $\epsilon/2$-far from each other, has distortion at most $\gamma(\epsilon)$.[7]

Moments of Distortion. For $1 \le q \le \infty$, define the *ℓ_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 ℓ_∞-distortion and the *average distortion* is expressed by the ℓ_1-distortion. The following was proved in [4].

Lemma 2.1. *([4]) Given a weighted graph $G = (V,E)$ on n vertices, if a spanner H has scaling distortion γ 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 3.1. (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 $\tilde{O}\left(\log^2 j\right)/\rho^2$.*

The main technical lemma is the following.

Lemma 3.1. *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 $O\left(\frac{\log^2 k}{\delta^2 \log\log k}\right)$.*

Before proving this lemma, we use it to prove Theorem 3.1.

[7]It can be verified that coarse scaling distortion γ implies scaling distortion γ.

Algorithm 1 `Greedy Spanner`$(G = (V,E), t)$

$H = (V, \emptyset)$.
for each edge $\{u,v\} \in E$, in non-decreasing order of weight, **do**
  if $d_H(u,v) > t \cdot w(u,v)$ **then**
    Add the edge $\{u,v\}$ to $E(H)$.
  end if
end for

Proof. (of Theorem 3.1) 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 3.1 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 $O\left(\frac{\log^2|K_i|}{\delta_i^2 \log\log|K_i|}\right) = O(2^{2i}\cdot i^3/\rho^2)$ 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 $2^{2i} \le 4\log^2 j$, and we conclude that

$$\begin{aligned} d_H(v_j,v_r) &\le d_{H_i}(v_j,v_r) \le O(2^{2i}\cdot i^3/\rho^2)\cdot d_G(v_j,v_r) \\ &\le \tilde{O}\left(\log^2 j\right)/\rho^2 \cdot d_G(v_j,v_r). \end{aligned}$$

as required. □

3.1 Proof of Lemma 3.1

We begin by devising a spanner with the required properties, but with distortion guarantee only for pairs in $K \times K$. Our construction is based on the simple greedy spanner algorithm (see Algorithm 1) on a modified version of the graph. Given a graph G and a parameter t, the greedy algorithm iterates over all the edges in increasing order of weights, and adds an edge to the spanner iff the current distortion (in the subgraph that consists of all the edges already added to the spanner) between its endpoints is greater than t. It can be verified that the spanner returned by the `Greedy Spanner` algorithm has distortion at most t, and also contains the MST of G (see [13]).

Let T be the unique MST of G. Let $G' = \left(V, T \cup \binom{K}{2}\right)$ be a graph that contains the MST, and in addition, the complete graph on the vertices of K. For $u, v \in K$

the edge $\{u,v\}$ will have weight $d_G(u,v)$ in G'. Note that each $e \in T$ is not the heaviest edge in any cycle of G', so T is the unique MST of G'. Let H' be the spanner returned by the `greedy spanner` algorithm on input G' with parameter $t = \frac{4\log^2 k}{\delta \log\log k}$. (Note that H' contains T.)

Let $\mathcal{E} = H' \setminus T$ and $s = \log k$. For any $i \geq 1$ let

$$E_i = \left\{ e \in \mathcal{E} : w(e) \in \left(s^{i-1}, s^i\right] \cdot \frac{\delta \cdot w(T)}{k^2} \right\} ,$$

and $E_0 = \mathcal{E} \setminus (\bigcup_{i \geq 1} E_i)$. The following lemma is essentially given in [18, Theorem 2] (obtained by picking $\epsilon = \frac{s}{\delta \log s}$ in that paper, and using that our distortion $t \geq (2s-1)(1+\epsilon)$).[8]

Lemma 3.2. ([18])

$$w(E_i) \leq O(w(T)) \cdot \frac{\delta \log s}{s^{1+(i-1)/s}} .$$

As $E_0 \subseteq \binom{K}{2}$, and every edge in E_0 has weight at most $\delta \cdot w(T)/k^2$, it holds that

$$w\left(E_0\right) \leq |E_0| \cdot \delta \cdot w(T)/k^2 \leq \delta \cdot w(T) .$$

By Lemma 3.2, the contribution of the other edges to the weight of $\mathcal{E}$ is at most

$$\begin{aligned} & \sum_{i=1}^{\infty} O(w(T)) \cdot \frac{\delta \log s}{s^{1+(i-1)/s}} \\ \leq \quad & O(w(T)) \cdot \frac{\delta \log s}{s} \sum_{i=0}^{\infty} e^{-(i \ln s)/s} \\ = \quad & O(w(T)) \cdot \frac{\delta \log s}{s} \cdot \frac{1}{1 - e^{-(\ln s)/s}} \\ = \quad & O(\delta \cdot w(T)) . \end{aligned}$$

We conclude that $\Psi\left(H'\right) \leq 1 + O(\delta)$. Observe that H' is not a subgraph of G, but this could easily be mended by replacing every edge in $\mathcal{E}$ with the corresponding shortest path in G. Clearly this will not increase the lightness of H'. Moreover, for any $u,v \in K$, it holds that

$$d_{H'}\left(u,v\right) \leq t \cdot d_{G'}\left(u,v\right) = t \cdot d_G\left(u,v\right) . \tag{3.1}$$

Observe that the spanner H' already implies the following result, obtained by choosing $K = V$. The dependence on δ is tight as shown in Section 6.

[8] Our definition of the E_i is slightly different than that of [18], we replaced n (recall that $n = |V|$) by k^2/δ. An inspection of their proof shows that such a change may be done, and results in having this factor in the lightness bound, instead of n.

Corollary 3.1. *For every $0 < \delta < 1$, every graph G on n vertices admits a spanner with (worst-case) distortion $O\left(\frac{\log^2 n}{\delta \log\log n}\right)$ and lightness $1+\delta$.*

It remains to extend H' so that every pair in $K \times V$ will suffer distortion at most $O(t/\delta)$. To this end, we use the following lemma which is a variation of shallow light trees [20, 8]. The proof can be found in Section 3.2.

Lemma 3.3. *Given a graph $G = (V,E)$, a parameter $\alpha > 1$, and a subset $K \subseteq V$, there exists a spanner S of G that **1)** contains the MST, **2)** has lightness $1 + \frac{2}{\alpha-1}$, and **3)** for any vertex $u \in V$, $d_S(u,K) \leq \alpha \cdot d_G(u,K)$.*

We proceed with the proof of Lemma 3.1. Let S be the spanner of Lemma 3.3 with respect to the set K and with $\alpha = 1 + 1/\delta$. Define our final spanner H as the union of H' and S. Note both H' and S have lightness $1+O(\delta)$ and contain the same MST, so H has lightness $1+O(\delta)$ as well. It remains to bound the distortion of a 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 3.3,

$$\begin{aligned} d_S(u,k_u) &= d_S(u,K) \leq \alpha \cdot d_G(u,K) \\ &\leq \alpha \cdot d_G(u,v) . \end{aligned} \tag{3.2}$$

Using the triangle inequality,

$$\begin{aligned} d_G(v,k_u) &\leq d_G(v,u) + d_G(u,k_u) \\ &\leq d_G(v,u) + d_S(u,k_u) \\ &\overset{(3.2)}{\leq} (\alpha+1) \cdot d_G(v,u) . \end{aligned} \tag{3.3}$$

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

$$\begin{aligned} d_{H'}(v,k_u) &\overset{(3.1)}{\leq} t \cdot d_G(v,k_u) \\ &\overset{(3.3)}{\leq} t\left(\alpha+1\right) \cdot d_G(v,u) . \end{aligned} \tag{3.4}$$

We conclude that

$$\begin{aligned} d_H\left(v,u\right) &\leq d_{H'}\left(v,k_u\right) + d_S\left(k_u,u\right) \\ &\overset{(3.2)\wedge(3.4)}{\leq} \left(t\left(\alpha+1\right)+\alpha\right) \cdot d_G(v,u) \\ &= O(t/\delta) \cdot d_G(v,u) , \end{aligned}$$

as required.

3.2 Proof of Lemma 3.3

The following lemma is implicit in [20].

Lemma 3.4. ([20]) *Given a weighted graph $G = (V, E, w)$, a parameter $\alpha > 1$ and a vertex $v \in V$, there exists a spanner H of G that **1)** contains the MST, **2)** has lightness $1 + \frac{2}{\alpha-1}$, and **3)** for any vertex $u \in V$, $d_H(u, v) \le \alpha \cdot d_G(u, v)$.*

The only difference between Lemma 3.3 and Lemma 3.4 is the replacement of the single vertex v by a set K. The standard way to handle multiple sources is to contract them to a single vertex, and grow the SLT from that vertex. Here one must be careful, since the MST of the resulting graph can be different than that of G.

Let $G' = \left(V, E \cup \binom{K}{2}\right)$ be a graph in which we add an edge of weight 0 between any two vertices in K (if such an edge already exists, reduce its weight to 0). We denote the new weights by w'. Let T be the unique MST of G, and let T' be an arbitrary MST of G'. Note that all the edges in $T' \setminus T$ must be from $\binom{K}{2}$, as any other edge is still the heaviest edge in a cycle of G'.

Fix any $v \in K$, and let H' be the spanner of Lemma 3.4 with respect to the graph G', the parameter α, and v. Let $\mathcal{E} = H' \setminus T'$. Note that the distance in T' between any two vertices in K is 0, so when bounding the weight of $\mathcal{E}$ we may assume that $\mathcal{E} \subseteq E \setminus \binom{K}{2}$. Hence for any edge $e \in \mathcal{E}$, $w(e) = w(e')$, which implies that $w'(\mathcal{E}) = w(\mathcal{E})$. Since $w'(T') + w'(\mathcal{E}) = w'(H') \le \left(1 + \frac{2}{\alpha-1}\right) \cdot w'(T')$, it holds that

$$w(\mathcal{E}) = w'(\mathcal{E}) \le \frac{2w'(T')}{\alpha - 1} \le \frac{2w(T)}{\alpha - 1} \ . \qquad (3.5)$$

Let $H = T \cup \mathcal{E}$. By (3.5) H has the desired lightness, and it remains to show the bound on the distortion from K. For any vertex $u \in V$, the shortest path in H' from u to K does not contain edges from $\binom{K}{2}$, and since all distances in H' among pairs in $\binom{K}{2}$ are 0, we have that

$$\begin{aligned} d_H(u, K) &= d_{H'}(u, K) = d_{H'}(u, v) \\ &\le \alpha \cdot d_{G'}(u, v) = \alpha \cdot d_G(u, K) \ , \end{aligned}$$

which concludes the proof of the lemma.

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 $f : X \to Y$,[9] 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 4.1. *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 α, 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 ϵ, 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 ϵ-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* ϵ-density-net *is a set $N \subseteq X$ such that: **1)** for all $v \in X$ there exists $u \in N$ with $d(v, u) \le 2R(v, \epsilon)$ and **2)** $|N| \le \frac{1}{\epsilon}$.*

Proof. (of Theorem 4.1) We begin by describing the desired priority ranking of X. For every integer $1 \le i \le \lceil \log n \rceil$ let $\epsilon_i = 2^{-i}$, and let $N_i \subseteq X$ be an ϵ_i-density-net in X. Take any priority ranking of X satisfying that every point $v \in N_i$ has priority at most $\left|\bigcup_{j=1}^{i} N_j\right| \le \sum_{j=1}^{i} |N_j|$. As for any j, $|N_j| \le \frac{1}{\epsilon_j} = 2^j$, each point in N_i has priority at most $\sum_{j=1}^{i} \frac{1}{\epsilon_j} \le \sum_{j=1}^{i} 2^j < 2^{i+1}$.

Let $f : X \to Y$ be some non-contractive embedding with prioritized distortion α with respect to the priorities we defined. Fix some $\epsilon \in (0, 1)$ and a pair $v, u \in V$ so that u is ϵ-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 ϵ-far from v, it holds that

$$d_X(v, v') \le 2R(v, \epsilon_i) \le 2R(v, \epsilon) \le 2d_X(v, u) \ . \qquad (4.6)$$

In particular, by the triangle inequality,

$$d_X(u, v') \le d_X(u, v) + d_X(v, v') \overset{(4.6)}{\le} 3d_X(u, v) \ . \qquad (4.7)$$

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

$$\begin{aligned} &d_Y(f(v), f(u)) \\ &\quad \le \quad d_Y(f(v), f(v')) + d_Y(f(v'), f(u)) \\ &\quad \le \quad \alpha(2^{i+1}) \cdot d_X(v, v') + \alpha(2^{i+1}) \cdot d_X(v', u) \\ &\overset{(4.6)\wedge(4.7)}{\le} \quad 5\alpha(2/\epsilon_i) \cdot d_X(v, u) \ . \end{aligned}$$

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

By the minimality of i it follows that $1/\epsilon_i \le 2/\epsilon$, and since α 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 3.1 and Theorem 4.1 we obtain the following.

THEOREM 4.2. *For any parameter $0 < \rho < 1$, any graph contains a spanner with coarse scaling distortion $\tilde{O}\left(\log^2(1/\epsilon)\right)/\rho^2$ and lightness $1+\rho$.*

Remark: By Lemma 2.1 it follows that this spanner has ℓ_q-distortion $\tilde{O}(q^2)/\rho^2$ 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 $\pi = v_1, \dots, v_n$. To achieve this, one may define a ranking which interleaves π with the ranking generated in the proof of Theorem 4.1. We leave the details to the reader.

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

THEOREM 4.3. *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 γ. Then, given a finite metric space (X, d_X) and a priority ranking $x_1, \dots, x_n$ 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 $x_1, \dots, x_n$ of the points of X, let $\delta = \min_{i\ne j} d_X(x_i, x_j)/2$. We define a new metric space (Z, d_Z) as follows. For every $1 \le i \le n$, 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 $i \ne j$, and $d_Z(u,v) = \delta$ otherwise. Observe that $|Z| = \sum_{i=1}^n |X_i| \le \sum_{i=1}^n (\mu(i)n + 1) \le 2n$.

We now use the embedding $f_Z : Z \to Y_Z$ with coarse scaling distortion γ, to define an embedding $f : X \to Y_Z$, by letting for every $1 \le i \le n$, $f(x_i) = f_Z(u_i)$ 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 $|X_i| \ge \mu(i)n \ge \frac{\mu(i)}{2}|Z|$, it holds that u_i, u_j are $\epsilon/2$-far from each other for $\epsilon = \mu(i)$. This implies that $\frac{d_{Y_Z}(f(x_i), f(x_j))}{d_X(x_i, x_j)} = \frac{d_{Y_Z}(f_Z(u_i), f_Z(u_j))}{d_Z(u_i, u_j)} \le \gamma(\mu(i))$. □

It follows from a result of [16] that the convergence condition on μ 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][10]. In particular, by applying a theorem of [3] we get the following:

THEOREM 4.4. *For every $1 \le p \le \infty$ and every finite metric space (X, d_X) and priority ranking of X, there exists an embedding with prioritized distortion $O(\log j)$ into $l_p^{O(\log |X|)}$.*

Remark: The proof of Theorem 4.1 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 the orginial 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 4.1 and Theorem 4.3, we infer that essentially any coarse scaling embedding can have such a one-sided guarantee, with a slightly worse dependence on ϵ, as claimed in the following corollary.

COROLLARY 4.1. *Fix a metric space (X, d) on n points. Let $\mathcal{Y}$ be a family of finite metric spaces as in Theorem 4.3. 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 $\sum_{i\ge 1}\mu(i) = 1$. .*

Proof. By the condition of Theorem 4.3, there exists $(Y, d_Y) \in \mathcal{Y}$ so that X embeds to Y with coarse scaling distortion $\gamma(\epsilon)$. According to Theorem 4.3, there is an embedding f with prioritized distortion $\gamma(\mu(i)))$ (w.r.t

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

to any fixed priority ranking π). We pick π to be the ordering required by Theorem 4.1, 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 bound on higher moments of the distortion.

THEOREM 5.1. *For any parameter $0 < \rho < 1$, any graph contains a spanning tree with scaling distortion $\tilde{O}(\sqrt{1/\epsilon})/\rho^2$ and lightness $1 + \rho$.*

It follows from Lemma 2.1 that the average distortion of the spanning tree obtained is $O(1/\rho^2)$. Moreover, the ℓ_q-distortion is $O(1/\rho^2)$ for any fixed $1 \le q < 2$, $\tilde{O}\left(\log^{2.5} n\right)/\rho^2$ for $q = 2$, and $\tilde{O}(n^{1-2/q})/\rho^2$ 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.[11]

LEMMA 5.1. *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 α (resp., β). 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].

[11] 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.

THEOREM 5.2. ([4]) *Any graph contains a spanning tree with scaling distortion $O(\sqrt{1/\epsilon})$.*

Now we can prove the main result.

Proof. (of Theorem 5.1) Let H be the spanner given by Theorem 4.2. Let T be a spanning tree of H constructed according to Theorem 5.2. By Lemma 5.1, T has scaling distortion $O(\sqrt{1/\epsilon}) \cdot \tilde{O}(\log^2(1/\epsilon))/\rho^2 = \tilde{O}(\sqrt{1/\epsilon})/\rho^2$ 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 $\mathbb{E}_Y\left[\frac{d_Y(f(u),f(v))}{d_X(u,v)}\right]$. The prioritized and scaling distortions are defined accordingly. We make use of the following result from [4].[12]

THEOREM 5.3. *([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 5.1 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 5.1, (while using Theorem 5.3 instead of Theorem 5.2), we obtain the following.

THEOREM 5.4. *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 $\tilde{O}(\log^4(1/\epsilon))/\rho^2$, such that every tree T in the support has lightness $1 + \rho$.*

It follows from Lemma 2.1 that the ℓ_q-distortion is $O(1/\rho^2)$, for every fixed $q \ge 1$.

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 ℓ_q-distortion as well). This shows that in the results of Theorem 5.1

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

and Theorem 4.2 the average distortion must depend on ρ.[13]

Lemma 6.1. *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 δ 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 $\bar{N} = V \setminus N$.

Consider $u \in \bar{N}$. By definition of N every weight 2 edge is further than δ steps away from u in H. It follows that there are at most $2\delta + 1$ vertices within distance at most δ 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

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

[13]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.

Finally,

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

□

7 Improvements and Future Directions

In a subsequent work, we improve the dependence of the distortion on ρ in Theorem 3.1 (and thus in Theorem 4.2, Theorem 5.1 and Theorem 5.4 as well), from $1/\rho^2$ to the optimal $1/\rho$. The main change in the construction is replacing Lemma 3.3 by the following.

Lemma 7.1. *Given a graph $G = (V, E)$, a subset $K \subseteq V$ of size k, and a parameter $\rho \geq 1$, there is a polynomial time algorithm that returns a spanner H of G which contains the MST, such that H has lightness $1 + \rho$ and*

$$\forall v \in V, \ d_H(v, k_v) \leq O\left(\log^2 k \ / \ \rho\right) \cdot d_G(v, k_v) .$$

(Recall that k_v is the closest vertex of K to v.)

Another, perhaps more ambitious, open problem rising from our work, would be to improve the bound in the prioritized distortion of Theorem 3.1 to $\tilde{O}(\log j/\rho)$. A possible approach would be to further improve the bound in Lemma 7.1 to $\tilde{O}(\log k/\rho)$, combined with an improved light spanner construction. In a very recent work (published in these proceedings), [14] construct a spanner with $(2t-1) \cdot (1+\epsilon)$ distortion and $O_\epsilon(n^{1/t})$ lightness. Choosing $t = \log n$ and $\epsilon = O(1)$, they obtain a spanner with distortion $O(\log n)$ and constant lightness. Thus, it may be possible to obtain a $1 + \rho$ lightness version of their result, combined with an appropriate improved version of Lemma 7.1.

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

[21] 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.