Spanners with Slack[⋆]

T-H. Hubert Chan[⋆⋆], Michael Dinitz[⋆⋆⋆], and Anupam Gupta

Carnegie Mellon University

Abstract. Given a metric (V, d), a *spanner* is a sparse graph whose shortest-path metric approximates the distance d to within a small multiplicative distortion. In this paper, we study the problem of *spanners with slack*: e.g., can we find sparse spanners where we are allowed to incur an arbitrarily large distortion on a small constant fraction of the distances, but are then required to incur only a constant (independent of n) distortion on the remaining distances? We answer this question in the affirmative, thus complementing similar recent results on embeddings with slack into ℓ_p spaces. For instance, we show that if we ignore an ϵ fraction of the distances, we can get spanners with $O(n)$ edges and $O(\log \frac{1}{\epsilon})$ distortion for the remaining distances.

We also show how to obtain sparse and low-weight spanners with slack from existing constructions of conventional spanners, and these techniques allow us to also obtain the best known results for distance oracles and distance labelings with slack. This paper complements similar results obtained in recent research on slack embeddings into normed metric spaces.

1 Introduction

The study of metric embeddings has been a central pursuit in algorithms research in the past decade: an embedding is a map from a metric space into a "simpler" metric space so that distances between points are changed by at most a small factor. More formally, given a *target class* $\mathcal{C}$ of metrics, an *embedding* of a finite metric space $M = (V, d)$ into the class $\mathcal{C}$ is a new metric space $M' = (V, d')$ such that $M' \in \mathcal{C}$. Most of the work on embeddings has used *distortion* as the fundamental measure of quality; the distortion of an embedding is the worst multiplicative factor by which distances are increased by the embedding[1]. Given the metric $M = (V, d)$ and the class $\mathcal{C}$, one natural goal is to find an embedding

⋆ Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213. This research was partly supported by the NSF CAREER award CCF-0448095, and by an Alfred P. Sloan Fellowship.

⋆⋆ Supported partly by a fellowship from the Croucher Foundation.

⋆⋆⋆ Supported partly by an NSF Graduate Research Fellowship and an ARCS Scholarship.

1 Formally, for an embedding $\varphi : (V, d) \to (V, d')$, the *distortion* is the smallest D so that $\exists\, \alpha, \beta \geq 1$ with $\alpha \cdot \beta \leq D$ such that $\frac{1}{\alpha}\, d(x, y) \leq d'(\varphi(x), \varphi(y)) \leq \beta\, d(x, y)$ for all pairs $x, y \in V \times V$. Note that this definition of distortion is no longer invariant under arbitrary scaling, since $\alpha, \beta \geq 1$; however, this is merely for notational convenience, and all our results can be cast in the usual definitions of distortion.

$\varphi((V,d)) = (V,d') \in \mathcal{C}$ such that the distortion of the map φ is minimized. Note that this notion of embedding is slightly non-standard but very natural, as it not only captures embeddings of metric spaces into geometric spaces (e.g., when $\mathcal{C}$ is the class of all Euclidean metrics, or the class of all ℓ_1 metrics), but also concepts such as *sparse spanners*, where the class $\mathcal{C}$ is the class of metrics generated by sparse graphs. In the rest of the paper, we will talk about metric embeddings in this general sense, thus including within its purview the results on spanners as well as those on embeddings into normed spaces, or for that matter, into distributions over tree metrics [5, 15]. (As an aside, let us note that the concept of distortion is often called "stretch" in the spanners literature, and we use the two terms interchangeably.)

In the theoretical community, the popularity of the notion of distortion/stretch has been driven by its applicability to approximation algorithms: if the embedding $\varphi : (V,d) \to (V,d')$ has a distortion of D, then the cost of solutions to some optimization problems on (V,d) and on (V,d') can only differ by some function of D; this idea has led to numerous approximation algorithms [20]. Seminal results in embeddings include the $O(\log n)$ distortion embeddings of arbitrary metrics into ℓ_p spaces [9], the fact that any metric admits an $O(\log n)$ stretch spanner with $O(n)$ edges [3], and that any metric can be embedded into a distribution of trees with distortion $O(\log n)$ [15]. (All the above three results are known to be tight.)

In parallel to this theoretical work, more applied communities have had much recent interest in embeddings (and more generally, but also somewhat vaguely, on problems of finding "simpler representations" of distance spaces). One example is the networking community, where there is much interest in taking the point-to-point latencies between nodes in a network, treating it as a metric space $M = (V,d)$ satisfying the triangle inequality,[2] and finding some simpler representation $M' = (V,d')$ of this resulting metric so that distances between nodes can be quickly and accurately computed in this "simpler" metric M'. (E.g., one may want to assign each node a short *label* so that the distance between two nodes can be inferred approximately by merely looking at their labels.)

Despite this similarity of interest, many of the theoretical results mentioned above have not been used widely in these applications; the logarithmic guarantees on the distortion are often deemed unacceptable. Indeed, the notion of distortion turns out to be a demanding and inflexible objective function, and the empirical works are often happy with guarantees of the following form: they allow some small fraction of the distances to be distorted by *arbitrary* amounts, but then seek very strong guarantees on the distortion incurred by the remaining large fraction of the distances. E.g., in the networking application above, we would be happy if *most* inter-node distances were correct and only a small fraction of distances would be estimated poorly. (This corresponds to some notion of being "good on average"; we will revisit this idea in Section 4.)

[2] While the triangle inequality can be violated by network latencies, empirical evidence suggests that these violations are small and/or infrequent enough to make metric methods a useful approach.

To remedy the situation, Kleinberg et. al. [21] defined the notion of *embeddings with slack*: in addition to the metric $M = (V, d)$ and the class $\mathcal{C}$ in the initial formulation above, we are also given a *slack parameter* ϵ. We now want to find a map $\varphi(M) = (V, d') \in \mathcal{C}$ whose distortion is bounded by some quantity $D(\epsilon)$ on all but an ϵ fraction of the pairs of points in $V \times V$. Note that we allow the distortion on the remaining ϵn^2 pairs of points to be arbitrarily large. The line of work starting with their paper, and furthered by Abraham et. al. [2] and [1] showed that very strong results were indeed possible: in fact, when allowed constant slack, one could get constant-distortion constant-dimensional embeddings!

Given these results for embeddings into normed spaces, it is natural to ask whether one can obtain similar results for spanners (and related constructs such as distance oracles and distance labelings). This paper studies spanners with slack, and gives strong guarantees that answer the question in the affirmative and complement the above results for embeddings into normed metric spaces.

1.1 Our Results

In this paper, we look at results on finding *spanners* when we are allowed to incur an arbitrary amount of distortion on an ϵ fraction of the distances. We say that $H = (V, E_H)$ is an *ϵ-slack spanner* of a metric (V, d) with distortion D if for each vertex $v \in V$, the furthest $(1-\epsilon)n$ vertices w from v satisfy $d(v, w) \leq d_H(v, w) \leq D\, d(v, w)$, i.e., the graph H maintains distances from each vertex to all but the closest ϵn vertices. Our first result is a general transformation procedure to convert standard constructions of spanners into spanners with slack.

Theorem 1 (General Conversion Theorem). *Suppose any metric admits an $\alpha(n)$-distortion spanner with $T(n)$ edges. Then given any metric (V, d) and any ϵ, we can find an ϵ-slack spanner H_ϵ for it with $O(\alpha(\frac{1}{\epsilon}))$-distortion and $n + T(\frac{1}{\epsilon})$ edges.*

Using this, it immediately follows that there are constant-slack spanners with linear number of edges and *constant* distortion!

Moreover, if we were given a graph $G = (V, E)$ whose shortest path metric is (V, d), we show in Section 3.1 how to extend the above theorem to output a *subgraph* of G that has slightly more edges. We can also extend Theorem 1 in another direction and find a spanner that has a small weight in addition to a small number of edges (see Section 3.2).

Note that in the above results, we are given an ϵ, and then output a spanner H_ϵ. We can do better, and output a *single* graph H such that it is an ϵ-slack spanner for *all ϵ simultaneously.* Two general conversion procedures are given in Section 4, which can be used to prove corollaries like:

Theorem 2 ("One Spanner for All Epsilons"). *Given any metric of size n, we can find a graph H with $O(n)$ edges that is an $O(\log \frac{1}{\epsilon})$-distortion ϵ-slack spanner for each ϵ. Moreover, if the metric is generated by a graph G, then H can be made a subgraph of G.*

The spanner H from the previous theorem preserves distances well on average. We consider two natural notions of average, which are defined in Section 2.

Theorem 3 ("Good on Average"). *The spanner H from Theorem 2 has $O(1)$ average distortion and $O(1)$ distortion-of-the-average, and moreover has $O(\log n)$ distortion in the worst case.*

We then turn our attention to the questions of constructing distance labelings and distance oracles, which are useful for resource-location applications mentioned in the introduction. Detailed results that appear in Section 5 include results like:

Theorem 4 (Labelings and Oracles). *Given any integer k, every metric admits ϵ-slack distance oracles where the query time and stretch are $O(k)$, and the space requirement is $O(n + k(\frac{1}{\epsilon})^{1+1/k})$ words. Moreover, there are ϵ-slack distance labeling schemes that uses $O((\frac{1}{\epsilon})^{1/k} \log^{1-1/k} \frac{1}{\epsilon})$ space and suffer distortion $O(k)$.*

1.2 Previous Work

This work builds on a large body of work on spanners dating back to the late 1980's [24, 3, 11, 4, 27] and still going strong [14, 8, 7, 6, 28]; see, e.g., [23] for many of the results. Spanners were initially studied for applications in network synchronization, but since then they have found myriad uses in network design and routing, as well as in many places where it is advisable to compactly store a graph without changing the distances much, such as in speeding up shortest path computations. Apart from the literature on finding spanners of general graphs, there has also been a large body of work on Euclidean spanners (see, e.g., [11, 4]), as well as work on spanners for doubling metrics [10, 19].

The study of distance labelings of graphs [22, 16, 25] requires assigning "short" labels to vertices so that the distance between two vertices can be inferred from their labels alone, without any additional information about the graph. It is known that if one wants to infer distances exactly, then one may require as many as n bits for each vertex; however, one can do with far less space if one just wants to estimate the distances approximately. A closely related concept is that of a *distance oracle*, which is a data structure that can be used to estimate distances between nodes using small space and fast query time. Exact distance oracles require one to store lots of information (e.g., the entire distance matrix) or large query time (to run a shortest-path computation), but fast and compact distance oracles for general graphs were given by Thorup and Zwick [27]; work on special classes of graphs appears in [26, 12, 18, 17].

As mentioned in the introduction, the work on embeddings with slack was initiated by Kleinberg et. al. [21], and many of the subsequent results were improved in Abraham et. al. [2] and [1]. Our techniques and results complement those in the two aforementioned papers. The notion of slack distance oracles has been studied in [1] and we extend some of the results marginally in Section 5.1.

Our work is closely related to the work on distance preservers [13] by Coppersmith and Elkin. A distance preserver is a subgraph that maintains distances *exactly* for a pre-specified subset of pairs of nodes, as opposed to our case in which the only guarantee is that a large fraction of pairs of nodes have their distances well-approximated. On the other hand, the work of Elkin and Peleg [14] shows that there exist sparse spanners such that for large distances that are at least logarithmically large, the multiplicative stretch can be close to 1. This agrees with our results in the sense that large distances can be better maintained than small distances.

2 Preliminaries and Notation

All metric spaces we consider in this paper are finite and the graphs we consider are undirected. Let (V, d) be a metric space, where $n = |V|$. The *ball* $B(x, r) = \{y \mid d(x, y) \leq r\}$ is the set of points at distance at most r from x. For $0 < \epsilon < 1$, $R(x, \epsilon)$ is the minimum distance r such that $|B(x, r)| \geq \epsilon n$. The point y is *ϵ-far* away from point x if $d(x, y) \geq R(x, \epsilon)$. Observe that all spanners H we consider are *non-contracting*; i.e., for any $x, y \in V$, $d(x, y) \leq d_H(x, y)$.

Definition 1 ((Uniform) Slack Spanner). *Given a metric (V, d) and $0 < \epsilon < 1$, a weighted graph $H = (V, E)$ with each edge $(u, v) \in E$ having weight $d(u, v)$ is an* α-spanner with ϵ-uniform slack *if for all $x, y \in V$ such that y is ϵ-far away from x, $d_H(x, y) \leq \alpha \cdot d(x, y)$. In general, α can be a function of ϵ and $|V|$. If the metric (V, d) is induced by some weighted graph G, we say that H is a* subgraph spanner *if H is a subgraph of G.*

In other words, an ϵ-uniform slack spanner is one such that for each point x, apart from the ϵn points closest to x, the distances from x to the rest of the points are well approximated. We call this concept "*uniform* slack" to be consistent with previous notation; all references to "ϵ-slack" in this paper mean "ϵ-uniform slack".[3]

Definition 2 (Gracefully degrading spanner). *A weighted graph H is an* $\alpha(\frac{1}{\epsilon})$-gracefully degrading spanner *for the metric (V, d) if for each $0 < \epsilon < 1$, H is an $\alpha(\frac{1}{\epsilon})$-spanner with ϵ-slack. The notion of subgraph spanner also applies analogously.*

We also consider two incomparable notions of "average" distortion; both have been considered previously in the literature, and we will construct spanners that are simultaneously good with respect to both these notions.

Definition 3 (Average Distortion). *The* average distortion *of a spanner H for a metric space (V, d) is* $\frac{1}{\binom{n}{2}} \sum_{\{x,y\} \in \binom{V}{2}} \frac{d_H(x,y)}{d(x,y)}$.

[3] For the record, there *is* a non-uniform notion of slack; see [2, Defn. 1.1] for details. Also, readers of [1] should note that ϵ-uniform slack embeddings are called "coarsely $(1 - \epsilon)$ partial embeddings" in that paper.

Definition 4 (Distortion of Averages). *The* distortion of averages *of a spanner H for a metric space (V,d) is* $\sum_{\{x,y\}\in\binom{V}{2}} d_H(x,y) / \sum_{\{x,y\}\in\binom{V}{2}} d(x,y)$.

Most of our algorithms make use of a small sample of points from the metric space V such that each point is "close" to some sample point:

Definition 5 (Density Net). *Given a metric space (V,d) with $n = |V|$, and $0 < \epsilon < 1$, an ϵ-*density net *is a set $N \subseteq V$ such that **(1)** for all $x \in V$, there exists $y \in N$ such that $d(x,y) \le 2R(x,\epsilon)$, and **(2)** $|N| \le \frac{1}{\epsilon}$.*

We will often refer to the nodes in N as *centers*. Note that the difference between an ϵ-net and an ϵ-density net is in the notion of "closeness"—here the allowed distance from x to its closest center depends on the density of points around x.

Lemma 1. *Given a metric space (V,d) and $0 < \epsilon < 1$, an ϵ-density net N can be found in polynomial time.*

Proof. For each point $x \in V$, let B_x denote the ball $B(x, R(x,\epsilon))$. We order the vertices in a list L by nondecreasing value of $R(\cdot,\epsilon)$, breaking ties arbitrarily, and initialize the set N to be empty. We remove the first vertex v from list L. If there exists $u \in N$ such that B_v intersects B_u, then we just discard v; otherwise, we add v to N and remove all vertices in the ball B_v from the list L. We repeat this process until the list L becomes empty and return N as our ϵ-density net.

We next show that the subset N returned satisfies the two properties given in Definition 5. Consider any point $x \in V$. We show that there is a point $y \in N$ within distance $2R(x,\epsilon)$ of x. If x is included in N, this is trivially true. Otherwise, either x was at some point the first vertex in list L and get discarded, or x was in some ball B_v and removed from list L. In the former case, there is some point $u \in N$ such that B_u intersects B_x. Since u appears before x in the initial list, $R(u,\epsilon) \le R(x,\epsilon)$ and hence the distance between x and the density-net point u is $d(u,x) \le R(u,\epsilon) + R(x,\epsilon) \le 2R(x,\epsilon)$. In the latter case, as v appear before x in the initial list, we also have $R(v,\epsilon) \le R(x,\epsilon)$ and so $d(x,v) \le R(v,\epsilon) \le R(x,\epsilon) \le 2R(x,\epsilon)$. To show that $|N| \le \frac{1}{\epsilon}$, observe that the intersection of B_x and B_y is empty for any two distinct points $x, y \in N$. Since for each $x \in N$, the ball B_x contains at least ϵn points, we conclude that $|N| \le \frac{1}{\epsilon}$.

3 Slack Spanners

In this section, we give a general transformation technique to convert $\alpha(n)$-spanners with $T(n)$ edges into ϵ-slack spanners with distortion $(5 + 6\alpha(\frac{1}{\epsilon}))$ and $n + T(\frac{1}{\epsilon})$ edges. Our construction is very simple:

> **Construction.** We first construct an ϵ-density net N as given in Lemma 1. Since $|N| \le \frac{1}{\epsilon}$, we can construct an $\alpha(\frac{1}{\epsilon})$-spanner $\widehat{H}$ for the set of centers N. Then, for each point $x \in X \setminus N$, we add an edge between x and its closest point in N to $\widehat{H}$; this gives us a spanner H for (V,d).

Theorem 5. *The spanner H has $n + T(\frac{1}{\epsilon})$ edges, and is a $(5 + 6\alpha(\frac{1}{\epsilon}))$-spanner with ϵ-uniform slack.*

Proof. First we bound the size of H. Since N has at most $\frac{1}{\epsilon}$ points, the spanner $\widehat{H}$ has at most $T(\frac{1}{\epsilon})$ edges. Moreover, for each point $x \in V \backslash N$, one extra edge is added. Hence, H has at most $n + T(\frac{1}{\epsilon})$ edges.

Next, we bound the stretch of H. Consider two points u and v such that v is ϵ-far away from u, i.e., $d(u,v) \geq R(u,\epsilon)$. Let u' be a closest point in N to which u is connected to in H (or set $u' = u$ if u is in N), and define v' similarly with respect to v. By the properties of the density net, the distance $d(u,u') \leq 2R(u,\epsilon) \leq 2d(u,v)$ and $d(v,v') \leq d(v,u') \leq d(v,u) + d(u,u') \leq 3d(u,v)$. Also, $d(u',v') \leq d(u',u) + d(u,v) + d(v,v') \leq 6d(u,v)$. This implies that

$$\begin{aligned} d_H(u,v) &\leq d(u,u') + d_H(u',v') + d(v',v) \leq 5d(u,v) + d_H(u',v') \\ &\leq 5d(u,v) + \alpha(\tfrac{1}{\epsilon})d(u',v') \leq 5d(u,v) + \alpha(\tfrac{1}{\epsilon})(6d(u,v)). \end{aligned}$$

As an example of how we apply Theorem 5, let us recall a well-known result about spanners for general metrics, from which we derive Corollary 1.

Theorem 6 (Spanners for general metrics [24, 3]). *For any metric of size n, there exists a $(2k-1)$-spanner with $O(n^{1+1/k})$ edges.*

Corollary 1 (Uniform slack spanners for general metrics). *For any metric, for any $0 < \epsilon < 1$, for any integer $k > 0$, there exists a $(12k-1)$-spanner with ϵ-uniform slack of size $n + O((\frac{1}{\epsilon})^{1+1/k})$.*

3.1 Subgraph spanners

Note that if the metric (V,d) was generated by a graph $G = (V,E)$, our previous construction may result in a spanner that is not a subgraph of the original graph G. We now give an alternative construction to obtain a subgraph spanner. Let us first recall a theorem by Coppersmith and Elkin [13] on subgraphs that preserve distances exactly for a given set P of pairs of vertices in a weighted graph $G = (V,E)$.

Theorem 7 ([13]). *Given a weighted graph $G = (V,E)$ and a set P of pairs of vertices, there exists a subgraph H of G with $O(n + \sqrt{n} \cdot |P|)$ edges such that for any $\{u,v\} \in P$, $d_H(u,v) = d_G(u,v)$.*

Construction of the Subgraph Spanner. As before, let N be an ϵ-density net, which we know has at most $\frac{1}{\epsilon}$ elements. We construct an $\alpha(\frac{1}{\epsilon})$-spanner H' of size $T(\frac{1}{\epsilon})$ on N, which we convert to a subgraph in the following manner. We take P to be the set of distinct pairs $\{u,v\}$ that are edges in H' to be the subgraph of G that preserves distances for all pairs in P in the manner as stated in Theorem 7. Finally, points in V are connected to N by shortest path trees rooted at the points in N, using edges in the given graph G.

Theorem 8 shows that the resulting subgraph spanner H contains a small number of edges and has small stretch. Applying the theorem to Theorem 6 gives a nice corollary.

Theorem 8. *The subgraph H is a $(5+6\alpha(\frac{1}{\epsilon}))$-spanner with ϵ-uniform slack and has $O(n+\sqrt{n}\cdot T(\frac{1}{\epsilon}))$ edges.*

Corollary 2 (Subgraph uniform slack spanners for general metrics). *For any metric, for any $0<\epsilon<1$, for any integer $k>0$, there exists a subgraph $(12k-1)$-spanner with ϵ-uniform slack of size $O(n+\sqrt{n}\cdot(\frac{1}{\epsilon})^{1+1/k})$.*

3.2 Low Weight Spanners

In some cases, we would like spanners which are not only sparse, but whose weight is also comparable to the weight of an MST on the metric (V,d). Due to lack of space, we merely mention a representative result here, and omit the details.

Proposition 1 (Low weight uniform slack spanner). *For any metric of size n, there exists an $O(\log\frac{1}{\epsilon})$-spanner with ϵ-uniform slack of size $O(n+\frac{1}{\epsilon})$ and weight $O(\log^2\frac{1}{\epsilon})$ times that of an MST.*

4 Gracefully Degrading Spanners and Notions of Average Distortion

In this section, we give general procedures to convert ordinary spanners into gracefully degrading spanners. Suppose we know how to construct ordinary $\alpha(n)$-spanners of size $T(n)$ for finite metrics of size n. Observe that typically, $\alpha(\cdot)$ is a sublinear function, such as $O(\log n)$. In particular, we assume that there exists $C,c>1$ such that $\alpha(n)\le C\alpha(n^{1/c})$.

> **Construction.** Take $\epsilon_0=n^{-1/c}$ (think of c as 2 or 4) and construct a 1-spanner H_0 for some ϵ_0-density net N_0 that has $O(n)$ edges. We also construct an $\alpha(n)$-spanner $\widehat{H}$ for the entire metric V. The gracefully degrading spanner consists of the union of $\widehat{H}$ and H_0, together with edges that connect each point in V to its closest point in N_0.

This simple construction gives us the following theorem on gracefully degrading spanners. Using similar techniques as in Section 3.1, we can obtain subgraph spanners as well. Applying Theorem 9 to Theorem 6 with $k=O(\log n)$, we obtain the result as promised in the introduction.

Theorem 9. *Suppose there exists an $\alpha(n)$-spanner of size $T(n)$ for any metric of size n, where $\alpha(\cdot)$ is a non-decreasing function such that there exists exists $C>1$ such that $\alpha(n)\le C\alpha(n^{1/2})$. Then, for any finite metric (V,d) of size n, there exists an $C\alpha(\frac{1}{\epsilon})$-gracefully degrading spanner of size at most $T(n)+O(n)$.*

If we have the stronger assumption that $\alpha(n)\le C\alpha(n^{1/4})$, then the gracefully degrading spanner can be made to be a subgraph of the weighted graph that induces the metric (V,d).

Corollary 3 (Gracefully degrading spanner for general metrics). *Any metric of size n has a $O(\log \frac{1}{\epsilon})$-gracefully degrading spanner H of size $O(n)$. If the metric is induced by some weighted graph G, then H can be made to be a subgraph of G.*

Since the greedy construction given in [24, 3] gives an $O(\log n)$-spanner with size $O(n)$ for any metric, we can show that any metric has a spanner that has $O(1)$ "average distortion" for both notions of average distortion given in Definitions 3 and 4 in the following theorem.

Theorem 10 ("Average Distortion"). *For any metric (V, d), there exists a spanner H with size $O(n)$ that has both constant average distortion and constant distortion of the average, and moreover has $O(\log n)$ stretch in the worst case. If the metric (V, d) is induced by some graph G, then H can be made to be a subgraph of G.*

5 Distance Oracles and Labelings

The techniques that we have developed for slack spanners also turn out to be useful for developing slack distance oracles and distance labelings. Distance oracles and labelings have been widely studied, and distance labelings were in fact one of the original motivations for the study of slack embeddings by Kleinberg et. al. [21]. Slack distance oracles and (implicitly) labelings were considered by Abraham et. al. [1], who gave both slack and gracefully degrading constructions. We give simple constructions with slightly better bounds for distance oracles, and give the first uniform slack labelings that do not use an embedding into ℓ_p, allowing us to bypass a lower bound from Abraham et al. [2].

5.1 Distance oracles

Thorup and Zwick [27] studied the problem of creating distance oracles for metric spaces. A distance oracle is a small data structure which allows fast queries for approximate distances. They give an oracle that, for any integer $k \geq 1$, takes $O(kn^{1+1/k})$ space, has $O(k)$ query time, and has stretch of $2k-1$. Slack distance oracles were first studied by Abraham et. al. [1], whose results we improve on for both uniform slack and gracefully degrading distance oracles. We first give a general transformational theorem. By using this transformation on the distance oracle of Thorup and Zwick [27, Theorem 3.1], we get a uniform slack distance oracle with the best known guarantee.

Theorem 11. *Suppose that there exists some distance oracle with $\alpha(n)$ stretch and $O(q(n))$ query time that uses $O(f(n))$ space. Then there exists a distance oracle with ϵ-uniform slack, $5 + 6\alpha(\frac{1}{\epsilon})$ stretch, and $O(q(\frac{1}{\epsilon}))$ query time that uses $O(n + f(\frac{1}{\epsilon}))$ space.*

Corollary 4 (Uniform slack distance oracle). *For every integer $k \geq 1$, there is a distance oracle with ϵ-uniform slack, $O(k)$ query time, and $12k-1$ stretch that uses $O(n + k(\frac{1}{\epsilon})^{1+1/k})$ space.*

The uniform slack distance oracle in the full version of Abraham et. al. [1] has stretch of only $6k-1$ and $O(k)$ query time, but uses $O(n \log n \log \frac{1}{\epsilon} + k \log n (\frac{1}{\epsilon} \log \frac{1}{\epsilon})^{1+1/k})$ space.

Gracefully degrading distance oracles The construction of gracefully degrading spanners in Section 4 can be easily modified to yield gracefully degrading distance oracles. Namely, we use two levels of distance oracles instead of two levels of spanners, where the oracle on the density net is exact. Combining this transformation theorem with the distance oracles of Thorup and Zwick [27] and the average case analysis of Theorem 10, we get the following:

Corollary 5 (Gracefully degrading distance oracle). *For any integer k with $1 \leq k \leq O(\log n)$, there is a distance oracle with worst case stretch of $2k-1$ and $O(k)$ query time that uses $O(kn^{1+1/k})$ space such that the average distortion and the distortion of average is $O(1)$.*

The gracefully degrading distance oracle of Abraham et. al. [1, Theorem 14] gives the same query time, worst case stretch, average distortion, and distortion of average. However, their construction modifies the standard Thorup and Zwick [27] construction by sampling the first level with probability $3n^{-1/k} \ln n$, and thus they use $O(n^{1+1/k} \log n)$ space. This is more than we use if $k = o(\log n)$ and the same if $k = \Theta(\log n)$.

5.2 Distance labels

A distance labeling is an assignment of labels to the vertices so that the approximate distance between any two vertices can be computed simply by looking at the two labels. The goals are to minimize the stretch, the size of the label, and the time needed to compute the distance given the two labels. We give the first uniform slack distance labeling that uses space independent of n. Note that any embedding of a metric into ℓ_p gives a distance labeling where the size of a label is the dimension of the embedding. Embeddings of this form were considered by Abraham et al. [2], who proved that the dimension must depend on $\log n$. Thus any distance labeling that uses a slack embedding into ℓ_p must use space that depends on $\log n$, whereas our labeling is independent of n.

Theorem 12. *Let (V, d) be a metric space with n points. Suppose that there exists a distance labeling where each label has size $O(f(n))$ and for any two points u, v it is possible to compute, in $O(q(n))$ time, an approximation to the distance between u and v with a stretch of at most $\alpha(n)$. Then there exists a distance labeling with ϵ-uniform slack such that every label has size $O(f(\frac{1}{\epsilon}))$, and computing distances up to a stretch of $5 + 6\alpha(\frac{1}{\epsilon})$ can be done in $O(q(\frac{1}{\epsilon}))$ time.*

We get the following corollary by simply applying Theorem 12 to the distance labeling of Thorup and Zwick [27, Theorem 3.4]. Note that the size of the labels is independent of n.

Corollary 6 (Uniform slack distance labeling). *Let (V, d) be a metric space on n points. Let $0 < \epsilon < 1$, and let k be an integer with $1 \leq k \leq \log \frac{1}{\epsilon}$. Then it is possible to assign each point a label that uses $O((\frac{1}{\epsilon})^{1/k} \log^{1-1/k} \frac{1}{\epsilon})$ space such that, given the labels of vertices u, v where v is ϵ-far from u, the distance $d(u, v)$ can be computed up to a stretch of $12k - 1$ in $O(k)$ time.*

We can also get gracefully degrading labelings. These labelings will be larger than the embeddings into ℓ_p given by Abraham et. al. [1, Theorem 10], but will have faster time complexity. We get the following result by combining a standard transformation theorem (which we omit) with the labelings of Thorup and Zwick [27] and Theorem 10.

Corollary 7 (Gracefully degrading distance labeling). *For any integer k with $1 \leq k \leq O(\log n)$, there is a distance labeling of any n point metric such that each label has size at most $O(n^{1/k} \log^{1-1/k} n)$, and given two labels it is possible to compute the distance between the two points up to a worst case stretch of $2k - 1$ in $O(k)$ time. Furthermore, the average distortion and the distortion of average are $O(1)$.*

References

1. I. Abraham, Y. Bartal, and O. Neiman. Advances in metric embedding theory. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 2006.
2. Ittai Abraham, Yair Bartal, T-H. Hubert Chan, Kedar Dhamdhere Dhamdhere, Anupam Gupta, Jon Kleinberg, Ofer Neiman, and Aleksandrs Slivkins. Metric embeddings with relaxed guarantees. In *FOCS '05: Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 83–100, Washington, DC, USA, 2005. IEEE Computer Society.
3. Ingo Althöfer, Gautam Das, David Dobkin, Deborah Joseph, and José Soares. On sparse spanners of weighted graphs. *Discrete Comput. Geom.*, 9(1), 1993.
4. Sunil Arya, Gautam Das, David M. Mount, Jeffrey S. Salowe, and Michiel H. M. Smid. Euclidean spanners: short, thin, and lanky. *STOC*, 1995.
5. Yair Bartal. Probabilistic approximations of metric spaces and its algorithmic applications. In *Proceedings of the37thAnnual IEEE Symposium on Foundations of Computer Science*, 1996.
6. Surender Baswana, Telikepalli Kavitha, Kurt Mehlhorn, and Seth Pettie. New constructions of (α, β)-spanners and purely additive spanners. In *SODA*, pages 672–681, 2005.
7. Surender Baswana and Sandeep Sen. Approximate distance oracles for unweighted graphs in õ(n^2) time. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 271–280, 2004.
8. Béla Bollobás, Don Coppersmith, and Michael Elkin. Sparse distance preservers and additive spanners (extended abstract). In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003)*, pages 414–423, New York, 2003. ACM.

9. Jean Bourgain. On Lipschitz embeddings of finite metric spaces in Hilbert space. *Israel Journal of Mathematics*, 52(1-2), 1985.
10. Hubert T-H. Chan, Anupam Gupta, Bruce M. Maggs, and Shuheng Zhou. On hierarchical routing in DOubling metrics. In *Proceedings of the16thAnnual ACM-SIAM Symposium on Discrete Algorithms*, 2005.
11. Barun Chandra, Gautam Das, Giri Narasimhan, and José Soares. New sparseness results on graph spanners. *Internat. J. Comput. Geom. Appl.*, 5(1-2), 1995. Eighth Annual ACM Symposium on Computational Geometry (Berlin, 1992).
12. S. Chaudhuri and C. D. Zaroliagis. Shortest paths in digraphs of small treewidth. I. Sequential algorithms. *Algorithmica*, 27(3-4):212–226, 2000. Treewidth.
13. Don Coppersmith and Michael Elkin. Sparse source-wise and pair-wise distance preservers. In *Proceedings of the16thAnnual ACM-SIAM Symposium on Discrete Algorithms*, 2005.
14. Michael Elkin and David Peleg. $(1+\epsilon,\beta)$-spanner constructions for general graphs. *SIAM J. Comput.*, 33(3):608–631 (electronic), 2004.
15. Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. In *Proceedings of theAnnual ACM Symposium on Theory of Computing*, 2003.
16. Cyril Gavoille, David Peleg, Stephane Perennes, and Ran Raz. Distance labeling in graphs. In *Proceedings of the12thAnnual ACM-SIAM Symposium on Discrete Algorithms*, 2001.
17. Joachim Gudmundsson, Christos Levcopoulos, Giri Narasimhan, and Michiel H. M. Smid. Approximate distance oracles for geometric graphs. In *SODA*, pages 828–837, 2002.
18. Anupam Gupta, Amit Kumar, and Rajeev Rastogi. Traveling with a Pez dispenser (or, routing issues in MPLS). *SIAM J. Comput.*, 34(2):453–474 (electronic), 2004/05.
19. Sariel Har-Peled and Manor Mendel. Fast construction of nets in low dimensional metrics, and their applications. *Symposium on Computational Geometry*, pages 150–158, 2005.
20. Piotr Indyk. Algorithmic aspects of geometric embeddings. In *Proceedings of the42ndAnnual IEEE Symposium on Foundations of Computer Science*, 2001.
21. Jon M. Kleinberg, Aleksandrs Slivkins, and Tom Wexler. Triangulation and embedding using small sets of beacons. In *Proceedings of the45thAnnual IEEE Symposium on Foundations of Computer Science*, 2004.
22. David Peleg. Proximity-preserving labeling schemes and thei applications. In *25th Workshop on Graph-Theoretic Concepts in Computer Science*, Lecture Notes in Computer Science 1665, 1999.
23. David Peleg. *Distributed computing*. Society for Industrial and Applied Mathematics (SIAM), 2000. A locality-sensitive approach.
24. David Peleg and Alejandro A. Schäffer. Graph spanners. *J. Graph Theory*, 13(1), 1989.
25. Kunal Talwar. Bypassing the embedding: Algorithms for low-dimensional metrics. In *Proceedings of the36thAnnual ACM Symposium on Theory of Computing*, 2004.
26. Mikkel Thorup. Compact oracles for reachability and approximate distances in planar digraphs. In *Proceedings of the42ndAnnual IEEE Symposium on Foundations of Computer Science*, 2001.
27. Mikkel Thorup and Uri Zwick. Approximate distance oracles. In *Proceedings of the33ndAnnual ACM Symposium on Theory of Computing*, 2001.
28. Mikkel Thorup and Uri Zwick. Spanners and emulators with sublinear distance errors. In *SODA*, pages 802–809, 2006.