Note on Bounded Degree Spanners for Doubling Metrics

T-H. Hubert Chan Anupam Gupta Bruce M. Maggs Shuheng Zhou

August 1, 2007

1 Introduction

This note gives a self-contained complete proof on the following result, which appears as Theorem 5.1 in the paper [CGMZ05].

Theorem 1.1 *Given a metric (V, d) with doubling dimension k, there exists a $(1+\epsilon)$-spanner such that the degree of every vertex is at most $(2+\frac{1}{\epsilon})^{O(k)}$.*

We focus on obtaining sparse representations of metrics: these are called *spanners*, and they have been studied extensively both for general and Euclidean metrics. Formally, a t-spanner for a metric $M = (V, d)$ is an undirected graph $G = (V, E)$ such that the distances according to d_G (the shortest-path metric of G) are close to the distances in d: i.e., $d(u, v) \leq d_G(u, v) \leq t\, d(u, v)$. Clearly, one can take a complete graph and obtain $t = 1$, and hence the quality of the spanner is typically measured by how few edges can G contain whilst maintaining a *stretch* of at most t. The notion of spanners has been widely studied for general metrics (see, e.g. [PS89, ADD$^+$93, CDNS95]), and for geometric distances (see, e.g., [CK95, Sal91, Vai91, ADM$^+$95]). Here, we are particularly interested in the case when the input metric has bounded doubling dimension and the spanner we want to construct has small stretch, i.e. $t = 1 + \varepsilon$, for small $\varepsilon > 0$. We show that for fixed ε and metrics with bounded doubling dimension, it is possible to construct linear sized $(1 + \varepsilon)$-spanners. Observe that any 1.5-spanner for a uniform metric on n points must be the complete graph. Hence, without any restriction on the input metric, it is not possible to construct an $(1 + \varepsilon)$-spanner with linear number of edges.

1.1 Notation and Preliminaries

We consider a finite metric $M = (V, d)$ where $|V| = n$. A metric has *doubling dimension* [GKL03] at most k if for every $R > 0$, every ball of radius R can be covered by 2^k balls of radius $R/2$.

Definition 1.2 (**$(1 + \varepsilon)$-spanner**) *Let (V, d) be a finite metric. Suppose $G = (V, E)$ is an undirected graph such that each edge $\{u, v\} \in E$ has weight $d(u, v)$, and $d_G(u, v)$ is the length of the shortest path between vertices u and v in G. The graph G, or equivalently, the set E of edges, is a* $(1+\varepsilon)$-spanner *for (V, d) if for all pairs u and v, $d_G(u, v)/d(u, v) \leq 1 + \varepsilon$.*

Definition 1.3 (Net) *Let S be a set of points in a metric (V, d), and $r > 0$. A subset N of S is an* r-net *for S if the following conditions hold.*

1. *For all $x \in S$, there exists some $y \in N$ such that $d(x, y) \leq r$.*
2. *For all $y \neq z \in N$, $d(y, z) > r$.*

The following fact states that for a doubling metric, you cannot pack too many points in some fixed ball such that the points are far away from one another.

Fact 1.4 *Suppose S is a set of points in a metric space with doubling dimension at most k. If S is contained in some ball of radius R and for all $y \neq z \in S$, $d(y,z) > r$, then $|S| \leq (4R/r)^k$.*

2 Construction of linear-sized $(1+\varepsilon)$-spanners

In this section, we show the existence of sparse spanners by giving an explicit construction. In particular, we have the following result.

Theorem 2.1 *Given a metric (V,d) with doubling dimension k, there exists a $(1+\varepsilon)$-spanner $\widehat{E}$ that has $(2+\frac{1}{\varepsilon})^{O(k)}n$ edges.*

The basic idea is to first construct a *net-tree* representing a sequence of nested nets of the metric space: this is fairly standard, and has been used earlier, e.g., in [Tal04, KL04, CGMZ05]. A nearly-linear-time construction of net-trees is given by Har-Peled and Mendel [HPM05].

Net trees are formally defined in the following.

Definition 2.2 (Hierarchical Tree) *A* hierarchical tree *for a set V is a pair (T,φ), where T is a rooted tree, and φ is a labeling function $\varphi : T \to V$ that labels each node of T with an element in V, such that the following conditions hold.*

1. *Every leaf is at the same depth from the root.*
2. *The function φ restricted to the leaves of T is a bijection into V.*
3. *If u is an internal node of T, then there exists a child v of u such that $\varphi(v) = \varphi(u)$. This implies that the nodes mapped by φ to any $x \in V$ form a connected subtree of T.*

Definition 2.3 (Net-Tree) *A* net tree *for a metric (V,d) is a hierarchical tree (T,φ) for the set V such that the following conditions hold.*

1. *Let N_i be the set of nodes of T that have height i. (The leaves have height 0.) Suppose δ is the minimum pairwise distance in (V,d). Let $0 < r_0 < \delta/2$, and $r_{i+1} = 2r_i$, for $i \geq 0$. (Hence, $r_i = 2^i r_0$.) Then, for $i \geq 0$, $\varphi(N_{i+1})$ is an r_{i+1}-net for $\varphi(N_i)$.*
2. *Let node $u \in N_i$, and its parent node be p_u. Then, $d(\varphi(u), \varphi(p_u)) \leq r_{i+1}$.*

In order to construct the spanner, we include an edge if the end points are from the same net in some scale and "reasonably close" to each other with respect to that scale. Using this idea, one can obtain the following theorem.

Theorem 2.4 *Given a finite metric $M = (V,d)$ with doubling dimension bounded by* $\dim$. *Let $\varepsilon > 0$ and (T,φ) be any net tree for M. For each $i \geq 0$, let*

$$E_i := \{\{u,v\} \mid u,v \in \varphi(N_i), d(u,v) \leq (4 + \tfrac{32}{\varepsilon}) \cdot r_i\} \setminus E_{i-1},$$

where E_{-1} is the empty set. (Here the parameters N_i, r_i are as in Definition 2.3.) Then $\widehat{E} := \cup_i E_i$ forms a $(1+\varepsilon)$-spanner for (V,d), with the number of edges being $|\widehat{E}| \leq (2+\frac{1}{\varepsilon})^{O(\dim)}|V|$.

We prove Theorem 2.4 through Lemmas 2.5 and 2.8.

Lemma 2.5 *The graph $(V,\widehat{E})$ is a $(1+\epsilon)$-spanner for (V,d).*

Proof: Let $\widehat{d}$ be the distance function induced by $(V,\widehat{E})$. Let $\gamma := 4 + \frac{32}{\varepsilon}$. We first show that each point in V is close to some point in $\varphi(N_i)$ under the metric $\widehat{d}$.

Claim 2.6 *For all $x \in V$, for all i, there exists $y \in \varphi(N_i)$ such that $\widehat{d}(x,y) \leq 2r_i$.*

Proof: We shall prove this by induction on i. For $i = 0$, $\varphi(N_0) = V$. Hence, the result holds trivially.

Suppose $i \geq 1$. By the induction hypothesis, there exists $y' \in \varphi(N_{i-1})$ such that $\widehat{d}(x,y') \leq 2r_{i-1}$. Since Y_i is a r_i-net of Y_{i-1}, there exists $y \in \varphi(N_i) \subseteq \varphi(N_{i-1})$ such that $d(y',y) \leq r_i = 2r_{i-1} \leq \gamma \cdot r_{i-1}$. Hence,, $(y',y) \in E_i \subseteq \widehat{E}$ and $\widehat{d}(y',y) = d(y',y)$, which is at most r_i.

Finally, by the triangle inequality, $\widehat{d}(x,y) \leq \widehat{d}(x,y') + \widehat{d}(y',y) \leq 2r_{i-1} + r_i = 2r_i$. ■

We next show that for any pair of vertices $x,y \in V$, $\widehat{d}(x,y) \leq (1+\epsilon)d(x,y)$. Suppose $r_i \leq d(x,y) < r_{i+1}$.

Suppose q is the integer such that $\frac{8}{2^q} \leq \epsilon < \frac{16}{2^q}$, i.e. $q := \lceil \log_2 \frac{8}{\epsilon} \rceil$.

We first consider the simple case when $i \leq q-1$. Then, $d(x,y) < 2^{i+1} r_0 \leq 2^q r_0 \leq \frac{16}{\epsilon} \cdot r_0 \leq \gamma \cdot r_0$. Since $x,y \in \varphi(N_0)$, it follows that $(x,y) \in \widehat{E}$ and $\widehat{d}(x,y) = d(x,y)$.

Next we consider the case when $i \geq q$. Let $j := i - q \geq 0$.

By Claim 2.6, there exist vertices $x', y' \in \varphi(N_j)$ such that $\widehat{d}(x,x') \leq 2r_{j+1}$ and $\widehat{d}(y,y') \leq 2r_{j+1}$.

We next show that $(x',y') \in \widehat{E}$. It suffices to show that $d(x',y') \leq \gamma \cdot r_j$.

$$\begin{array}{rcll}
d(x',y') & \leq & d(x',x) + d(x,y) + d(y,y') & \text{(Triangle inequality)} \\
& \leq & 2r_j + r_{i+1} + 2r_j & \text{(Choice of } x', y' \text{ and } i) \\
& = & r_j(4 + 2 \cdot 2^q) & (i = j + q) \\
& \leq & r_j(4 + \frac{32}{\epsilon}) & (2^q < \frac{16}{\epsilon}) \\
& = & \gamma \cdot r_j &
\end{array}$$

Hence, we have $\widehat{d}(x',y') = d(x',y')$. Note that by the triangle inequality,

$$d(x',y') \leq d(x',x) + d(x,y) + d(y,y') \leq 4 \cdot r_j + d(x,y). \tag{2.1}$$

Finally, we obtain the desired upper bound for $\widehat{d}(x,y)$.

$$\begin{array}{rcll}
\widehat{d}(x,y) & \leq & \widehat{d}(x,x') + \widehat{d}(x',y') + \widehat{d}(y',y) & \text{(Triangle inequality)} \\
& \leq & 8 \cdot r_j + d(x,y) & \text{(Choice of } x', y' \text{ and (2.1))} \\
& = & \frac{8}{2^q} \cdot r_i + d(x,y) & (j = i - q) \\
& \leq & (1 + \frac{8}{2^q}) d(x,y) & (r_i \leq d(x,y)) \\
& \leq & (1+\epsilon) d(x,y) & (\frac{8}{2^q} \leq \epsilon)
\end{array}$$

■

We next proceed to show that the spanner $(V, \widehat{E})$ is sparse. We first show that for each vertex u, for each i, the number of edges in E_i incident on u is small.

Claim 2.7 *Define $\Gamma_i(u) := \{v \in V : \{u,v\} \in E_i\}$. Then, $|\Gamma_i(u)| \leq (4\gamma)^k$.*

Proof: Observe that $\Gamma_i(u)$ is contained in a ball of radius at most $\gamma \cdot r_i$ centered at u. Moreover, since $S \subseteq \varphi(N_i)$, any two points in S must be more than r_i apart. Hence, from Fact 1.4, it follow that $|\Gamma_i(u)| \leq (4\gamma)^k$. ■

Lemma 2.8 *The number of edges in $\widehat{E}$ is at most $(2 + \frac{1}{\epsilon})^{O(k)} n$.*

Proof: It suffice to show that the edges of $\widehat{E}$ can be directed such that each vertex has out-degree bounded by $(2+\frac{1}{\epsilon})^{O(k)}$.

For each $v \in V$, define $i^*(v) := \max\{i \,|\, v \in \varphi(N_i)\}$. For each edge $(u,v) \in \widehat{E}$, we direct the edge from u to v if $i^*(u) < i^*(v)$. If $i^*(u) = i^*(v)$, the edge can be directed arbitrarily. By *arc* (u,v), we mean an edge that is directed from vertex u to vertex v.

We now bound the out-degree of vertex u. Suppose there exists an arc $(u,v) \in E_i$.

By definition of E_i, $d(u,v) \leq \gamma \cdot r_i$. Recall $p = \lceil \log_2 \gamma \rceil$. Hence, it is not possible for both u and v to be contained in Y_{i+p}. Since $i^*(u) \leq i^*(v)$, it follows that $i^*(u) \leq i+p$. On the other hand, $u \in Y_i$ and so $i^*(u) \geq i$. So, $i^*(u) - p \leq i \leq i^*(u)$.

There are at most $p+1 = O(\log \gamma)$ values of i such that E_i contains an edge directed out of u. By Claim 2.7, for each i, the number of edges in E_i incident on u is at most $(4\gamma)^k$.

Hence, the total number of edges in $\widehat{E}$ directed out of u is $(4\gamma)^k \cdot O(\log \gamma) = (2+\frac{1}{\epsilon})^{O(k)}$.

■

3 Construction of bounded-degree $(1+\varepsilon)$-spanners

We have shown that the edges in $\widehat{E}$ can be directed such that the out-degree of every vertex is bounded. We next describe how to modify $\widehat{E}$ to get another set of edges $\tilde{E}$ that has size at most that of $\widehat{E}$, but the resulting undirected graph $(V,\tilde{E})$ has bounded degree (Lemma 3.1). Moreover, we show in Lemma 3.2 that the modification preserves distances between vertices.

We form the new graph $(V,\tilde{E})$ by modifying the directed graph $(V,\widehat{E})$ in the following way.

> **Modification Procedure.** Let l be the smallest positive integer such that $\frac{1}{2^{l-1}} \leq \epsilon$. Then, $l = O(\log \frac{1}{\epsilon})$.
>
> For each i and point u, define $M_i(u)$ to be the set of vertices w such that $w \in \Gamma_i(u)$ and (w,u) is directed into u in $\widehat{E}$.
>
> Let $I_u := \{i \,|\, \exists v \in M_i(u)\}$. Suppose the elements of I_u are listed in increasing order $i_1 < i_2 < \cdots$. To avoid double subscripts, we write $M^u_j := M_{i_j}(u)$.
>
> We next modify arcs going into each vertex u in the following manner. For $1 \leq j \leq l$, we keep the arcs directed from M^u_j to u. For $j > l$, we pick an arbitrary vertex $w \in M^u_{j-l}$ and for each point $v \in M^u_j$, replace the arc (v,u) by the arc (v,w).
>
> Observe that since M^u_j is defined with respect to the directed graph $(V,\hat{E})$, the ordering of the u's for which the modification is carried out is not important.

Let $(V,\tilde{E})$ be the resulting undirected graph. Since every edge in $\widehat{E}$ is either kept or replaced by another edge (which might be already in $\widehat{E}$), $|\tilde{E}| \leq |\widehat{E}|$.

Lemma 3.1 *Every vertex in $(V,\tilde{E})$ has degree bounded by $(2+\frac{1}{\epsilon})^{O(k)}$.*

Proof: Let α be an upper bound for the out-degree of the graph $(V,\widehat{E})$. From Lemma 2.8, we have $\alpha = (2+\frac{1}{\epsilon})^{O(k)}$. Let β be an upper bound for $|M_i(u)|$. We have $\beta \leq |\Gamma_i(u)| = (2+\frac{1}{\epsilon})^{O(k)}$.

We next bound the maximum degree of a vertex in $(V,\tilde{E})$. Consider a vertex $u \in V$. The edges incident on u can be grouped as follows.

1. There are at most α edges directed out of u in $\widehat{E}$.

2. Out of the edges in $\widehat{E}$ directed into u, at most βl remain in $\tilde{E}$.

3. New edges can be attached to u in $(V, \tilde{E})$. For each arc (u, v) directed out of u in $\widehat{E}$, there can be at most β new edges attaching to u in $\tilde{E}$. The reason is (u, v) can be in exactly one E_i and so there exists unique j such that $u \in M_j^v$. Hence, there could be potentially only at most $|M_{j+l}^v|$ new arcs directed into u because of the arc (u, v) in $\widehat{E}$.

Hence, the number of edges incident on u in $(V, \tilde{E})$ is bounded by $\alpha + \beta l + \alpha\beta = (2 + \frac{1}{\epsilon})^{O(k)}$.

■

We next show that the modification from $(V, \widehat{E})$ to $(V, \tilde{E})$ does not increase the distance between any pair of vertices too much.

Lemma 3.2 *Suppose $\tilde{d}$ is the metric induced by $(V, \tilde{E})$. Then, $\tilde{d} \leq (1 + 4\epsilon)\widehat{d}$.*

Proof: It suffices to show that for each edge $(v, u) \in \widehat{E}$ removed, $\tilde{d}(v, u) \leq (1 + 4\epsilon)d(v, u)$.

Suppose (v, u) in $\widehat{E}$ is directed into u . Then, by construction, $v \in M_j^u$ for some $j > l$.

Let $v_0 = v$. Then, from our construction, for $0 \leq s \leq s_j := \left\lfloor \frac{j-1}{l} \right\rfloor$, there exists $v_s \in M_{j-sl}^u$ such that for $0 \leq s < s_j$, $(v_s, v_{s+1}) \in \tilde{E}$, and $(v_{s_j}, u) \in \tilde{E}$. Then, there is a path in $(V, \tilde{E})$ going from v to u traversing vertices in the following order: $v = v_0, v_1, \ldots, v_{s_j}, u$. By the triangle inequality, the quantity $\tilde{d}(v, u)$ is at most the length of this path, which we show is comparable to $d(v, u)$.

Claim 3.3 *For $0 \leq s < s_j$, $d(u, v_{s+1}) \leq \epsilon d(u, v_s)$.*

Proof: Note that $v_{s+1} \in M_i(u)$ and $v_s \in M_j(u)$ for some i and j. From step 3 of our construction, $j - i \geq l$. Since $d(v_s, u) \geq \gamma \cdot r_{j-1}$ and $d(v_{s+1}, u) \leq \gamma \cdot r_i$, it follows that $d(v_{s+1}, u) \leq \frac{2}{2^l} d(v_s, u) \leq \epsilon d(v_s, u)$.

■

Claim 3.4 *For $0 \leq s \leq s_j$, $d(v_s, u) \leq \epsilon^s d(v_0, u)$.*

Proof: The claim can be proved by induction on s and using Claim 3.3. ■

From the triangle inequality and Claims 3.3 and 3.4, we have

$$d(v_s, v_{s+1}) \leq d(v_s, u) + d(u, v_{s+1}) \leq (1 + \epsilon)d(v_s, u) \leq (1 + \epsilon)\epsilon^s d(v_0, u) \tag{3.2}$$

Finally, we have

$$\begin{array}{rcll} \tilde{d}(v, u) & \leq & \sum_{r=0}^{s_j-1} d(v_s, v_{s+1}) + d(v_{s_j}, u) & \text{(Triangle inequality)} \\ & \leq & \sum_{s=0}^{s_j-1} (1 + \epsilon)\epsilon^s d(v_0, u) + \epsilon^{s_j} d(v_0, u) & \text{((3.2) and Claim 3.4)} \\ & \leq & \frac{1+\epsilon}{1-\epsilon} d(v_0, u) & \\ & \leq & (1 + 4\epsilon)d(v, u) & \end{array}$$

The last inequality follows from the fact that for $0 < \epsilon < \frac{1}{2}$, $\frac{1+\epsilon}{1-\epsilon} \leq 1 + 4\epsilon$. ■

Finally, we show that $(V, \tilde{E})$ is the desired spanner.

Theorem 3.5 *Given a metric (V, d) with doubling dimension k, there exists a $(1+\epsilon)$-spanner such that the degree of every vertex is at most $(2+\frac{1}{\epsilon})^{O(k)}$.*

Proof: We show that $\tilde{E}$ gives the desired spanner. Lemma 3.1 gives the bound on its degree. From Lemmas 2.5 and 3.2, we have $\tilde{d} \leq (1+4\epsilon)\widehat{d} \leq (1+4\epsilon)(1+\epsilon)d \leq (1+7\epsilon)d$, for $0 < \epsilon \leq \frac{1}{2}$. Substituting $\epsilon := \frac{\epsilon'}{7}$ gives the required result.

■

References

[ADD+93] Ingo Althöfer, Gautam Das, David Dobkin, Deborah Joseph, and José Soares. On sparse spanners of weighted graphs. *Discrete Comput. Geom.*, 9(1):81–100, 1993.

[ADM+95] Sunil Arya, Gautam Das, David M. Mount, Jeffrey S. Salowe, and Michiel H. M. Smid. Euclidean spanners: short, thin, and lanky. *STOC*, pages 489–498, 1995.

[CDNS95] Barun Chandra, Gautam Das, Giri Narasimhan, and José Soares. New sparseness results on graph spanners. *Internat. J. Comput. Geom. Appl.*, 5(1-2):125–144, 1995. Eighth Annual ACM Symposium on Computational Geometry (Berlin, 1992).

[CGMZ05] Hubert T-H. Chan, Anupam Gupta, Bruce M. Maggs, and Shuheng Zhou. On hierarchical routing in doubling metrics. *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 762–771, 2005.

[CK95] Paul B. Callahan and S. Rao Kosaraju. A decomposition of multidimensional point sets with applications to k-nearest-neighbors and n-body potential fields. *J. Assoc. Comput. Mach.*, 42(1):67–90, 1995.

[GKL03] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low–distortion embeddings. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 534–543, 2003.

[HPM05] 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.

[KL04] Robert Krauthgamer and James R. Lee. Navigating nets: Simple algorithms for proximity search. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 791–801, 2004.

[PS89] David Peleg and Alejandro A. Schäffer. Graph spanners. In *J. Graph Theory 13*, pages 99–116, 1989.

[Sal91] Jeffrey S. Salowe. Constructing multidimensional spanner graphs. In *Internat. J. Comput. Geom. Appl. 1*, pages 99–107, 1991.

[Tal04] Kunal Talwar. Bypassing the embedding: Algorithms for low-dimensional metrics. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 281–290, 2004.

[Vai91] Pravin M. Vaidya. A sparse graph almost as good as the complete graph on points in k dimensions. In *Discrete Comput. Geom. 6*, pages 369–381, 1991.