On some combinatorial problems in metric spaces of bounded doubling dimension

Michiel Smid*

February 5, 2010

Abstract

A metric space has doubling dimension d if for every $\rho > 0$, every ball of radius ρ can be covered by at most 2^d balls of radius $\rho/2$. This generalizes the Euclidean dimension, because the doubling dimension of Euclidean space $\mathbb{R}^d$ is proportional to d. The following results are shown, for any $d \geq 1$ and any metric space of size n and doubling dimension d: First, the maximum number of diametral pairs is $\Theta(n^2)$. Second, if $d = 1$, the maximum possible weights of the minimum spanning tree and the all-nearest neighbors graph are $\Theta(R \log n)$ and $\Theta(R)$, respectively, where R is the minimum radius of any ball containing all elements of the metric space. Finally, if $d > 1$, the maximum possible weights of both the minimum spanning tree and the all-nearest neighbors graph are $\Theta(Rn^{1-1/d})$. These results show that, for $1 \leq d \leq 3$, metric spaces of doubling dimension d behave differently than their Euclidean counterparts.

1 Introduction

Let S be a set of n points in $\mathbb{R}^d$, where $d \geq 1$ is a constant. Let G be a graph with vertex set S, in which the weight of any edge (p, q) is defined to be the Euclidean distance $|pq|$ between p and q. The weight $wt(G)$ of G is defined to be the sum of the weights of the edges in G.

*School of Computer Science, Carleton University, Ottawa, Canada. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

We are interested in the maximum values (over all sets of n points) of the number of diametral pairs, the weight of the minimum spanning tree, and the weight of the all-nearest neighbors graph.

Diametral pairs: The diameter of S is defined to be the maximum distance between any pair of points in S. Two points p and q in S form a *diametral pair*, if $|pq|$ is equal to the diameter of S. Let $DP(S)$ be the number of diametral pairs in the set S. Obviously, $DP(S) = 1$ if $d = 1$. If $d = 2$, the maximum possible value of $DP(S)$ is n, whereas if $d = 3$, the maximum possible value is $2n - 2$. If $d \geq 4$, the value of $DP(S)$ can be as large as $\Theta(n^2)$. For proofs of these claims, see Chapter 13 in Pach and Agarwal [5].

Minimum spanning tree: If S is contained in the d-dimensional cube $[0, L]^d$, then the weight $wt(MST(S))$ of its minimum spanning tree $MST(S)$ is $O(n^{1-1/d}L)$. This result was first shown by Few [2]. Subsequently, Smith [7], and Steele and Snyder [8] gave tighter analyses in terms of the constant factor in the Big-Oh bound. For the set S consisting of the n vertices on an $n^{1/d} \times \ldots \times n^{1/d}$ grid with cells having sides of length $n^{-1/d}L$, we have $wt(MST(S)) = \Omega(n^{1-1/d}L)$. Thus, the above upper bound on $wt(MST(S))$ is tight.

All-nearest neighbors graph: For any point p in S, let $NN_S(p)$ denote a nearest neighbor of p in S. That is, $NN_S(p)$ is a point q in $S \setminus \{p\}$ for which the Euclidean distance $|pq|$ is minimum. The all-nearest neighbors graph $ANN(S)$ is defined to be the directed graph with vertex set S and edge set $\{(p, NN_S(p)) : p \in S\}$. Assume again that the point set S is contained in the cube $[0, L]^d$. Since $ANN(S)$ is a subgraph of $MST(S)$, we have $wt(ANN(S)) = O(n^{1-1/d}L)$. The grid example given above shows that this upper bound is tight.

All results mentioned above are valid in the d-dimensional Euclidean space. A natural question to ask is whether or not similar results hold in an arbitrary metric space.

Let S be a finite set of elements, called points, that has a distance function associated with it. This function assigns to any two points p and q in S a real number $|pq|$, which is called the distance between p and q. The set S is called a *metric space*, if for all points p, q, and r in S,

1. $|pp| = 0$,
2. if $p \neq q$, then $|pq| > 0$,
3. $|pq| = |qp|$, and
4. $|pq| \leq |pr| + |rq|$.

The fourth property is called the *triangle inequality.*

If c is a point in S and $R > 0$ is a real number, then the *ball* with center c and radius R is defined to be the set $\{p \in S : |cp| \leq R\}$. The *radius* of a subset S' of S is defined to be the minimum radius of any ball that contains S'. Observe that the center of this ball is not necessarily an element of S'.

For example, the set S of n points, in which each pair of distinct points has distance R, is a metric space of radius R. Obviously, we have $DP(S) = \binom{n}{2}$, $wt(MST(S)) = (n-1)R$, and $wt(ANN(S)) = nR$. Moreover, these are the largest possible values in terms of n and R.

Thus, in order to get non-trivial results, we have to consider restricted classes of metric spaces. In this paper, we consider metric spaces of bounded doubling dimension, which were introduced by Assouad [1]; see also the book by Heinonen [4].

Definition 1 *Let S be a finite metric space and let λ be the smallest integer such that the following is true: For each real number $\rho > 0$, every ball in S of radius ρ can be covered by at most λ balls of radius $\rho/2$. The* doubling dimension *of the metric space S is defined to be* $\log \lambda$.

Observe that every metric space of n points has doubling dimension at most $\log n$. Furthermore, it is not difficult to prove that the doubling dimension of the Euclidean metric in $\mathbb{R}^d$ is proportional to d.

In this paper, we prove the following results, for any metric space S consisting of n points and having doubling dimension d:

1. If $d \geq 1$, then $DP(S)$ is $\Theta(n^2)$ in the worst case. Thus, the maximum possible value of $DP(S)$ is very different from the corresponding value in Euclidean space $\mathbb{R}^d$ for $1 \leq d \leq 3$. In fact, we show that, for $d = 1$, the maximum possible value of $DP(S)$ is $n^2/4$. For $d \geq 4$, the bound coincides with the corresponding bound for $\mathbb{R}^d$.

2. If $d = 1$, then $wt(MST(S))$ is $\Theta(R \log n)$ in the worst case, whereas $wt(ANN(S))$ is $\Theta(R)$ in the worst case. Thus, metric spaces of doubling dimension 1 behave differently than their one-dimensional Euclidean counterparts.

3. If $d > 1$, then both $wt(MST(S))$ and $wt(ANN(S))$ are $\Theta(Rn^{1-1/d})$ in the worst case. Thus, these bounds coincide with the corresponding bounds for Euclidean space $\mathbb{R}^d$.

Related work: In recent years, metric spaces of bounded doubling dimension have received a lot of attention in the algorithms literature; see Har-Peled and Mendel [3] and the references given in that paper. In [3], it was already observed that these spaces are different from their Euclidean counterparts: In Euclidean space $\mathbb{R}^d$, the all-nearest neighbors graph can be computed in $O(n \log n)$ time, whereas the worst-case running time becomes $\Theta(n^2)$ even if the doubling dimension is equal to 1.

Smid [6] defined the following metric space, based on an earlier example in [3]: Let $S = \{p_1, p_2, \dots, p_n\}$ and let $0 < \epsilon < 1$ be a real number. For each i and j with $1 \leq i \leq j \leq n$, we define

$$|p_i p_j| = |p_j p_i| = \begin{cases} 0 & \text{if } i = j, \\ 4^j & \text{if } i = 1 \text{ and } j > 1, \\ 4^j + \epsilon & \text{otherwise.} \end{cases}$$

Then S is a metric space of doubling dimension 1 and radius $R = 4^n + \epsilon$. We have $NN_S(p_1) = p_2$ and, for each j with $2 \leq j \leq n$, $NN_S(p_j) = p_1$. Thus, both $ANN(S)$ and $MST(S)$ are equal to the star-graph consisting of all edges (p_1, p_j), $2 \leq j \leq n$. (In contrast, it is well known that in Euclidean space $\mathbb{R}^d$, the maximum degrees of both these graphs are upper bounded by a function that depends only on d.) Both $wt(ANN(S))$ and $wt(MST(S))$ are close to $4R/3$.

The rest of this paper is organized as follows. In Section 2, we prove the upper bounds on $DP(S)$, $wt(ANN(S))$, and $wt(MST(S))$, for any finite metric space of doubling dimension d. The proofs will be by induction on the number of points in S. Some care has to be taken, because the doubling dimension d' of a subset of S may be larger than d. We show that d' is always at most $2d$.

In Section 3, we prove that the upper bounds in Section 2 are tight. The lower bounds for $DP(S)$, $wt(ANN(S))$, and $wt(MST(S))$ are in fact obtained by one single class of metric spaces.

We conclude in Section 4 by presenting some more results that show that metric spaces of doubling dimension 1 are very different than the Euclidean space $\mathbb{R}^1$.

2 The upper bounds

We start by showing that $DP(S) \leq n^2/4$ for every metric space S of size n and doubling dimension 1. The proof uses the following lemma which states that in any isosceles triangle, the base is shorter than the two sides of equal length.

Lemma 1 *Let S be a finite metric space of doubling dimension 1, and let p, q, and r be three pairwise distinct points in S such that $|pq| = |pr|$. Then $|qr| < |pq|$.*

Proof. The proof is by contradiction. Let p, q, and r be a "smallest" triple of points for which the lemma does not hold. That is, let ℓ be the smallest real number such that there exist three points p, q, and r in S such that $\ell = |pq| = |pr|$ and $|qr| \geq \ell$.

Let B be the ball with center p and radius ℓ. Observe that B contains the points p, q, and r. Since S has doubling dimension 1, we can cover B by two balls B_1 and B_2 centered at c_1 and c_2, respectively, and having radius $\ell/2$. We may assume without loss of generality that p is contained in B_1.

Let $\ell' = |qr|$. If both q and r are contained in B_2, then

$$\ell' = |qr| \leq |qc_2| + |c_2r| \leq \ell/2 + \ell/2 = \ell \leq \ell'.$$

It follows that $|c_2q| = |c_2r| = \ell/2$. Since $|qr| = \ell \geq \ell/2$, the triple c_2, q, and r forms a "smaller" counterexample to the lemma, which is a contradiction.

Thus, q and r are not both contained in B_2. We may assume without loss of generality that q is contained in B_1. We have

$$\ell = |pq| \leq |pc_1| + |c_1q| \leq \ell/2 + \ell/2 = \ell,$$

which implies that $|c_1p| = |c_1q| = \ell/2$. Since $|pq| = \ell \geq \ell/2$, the triple c_1, p, and q forms a "smaller" counterexample to the lemma, which is again a contradiction. ∎

Theorem 1 *Let S be a metric space of doubling dimension* 1 *and consisting of n points. Then $DP(S) \leq n^2/4$.*

Proof. Let R be the diameter of S, let c be a point in S, and let B be the ball with center c and radius R. Observe that B contains all points of S. Cover B by two balls B_1 and B_2, both having radius $R/2$, and define $S_1 = S \cap B_1$ and $S_2 = S \setminus S_1$.

Assume that S_1 contains two points p and q such that $|pq| = R$. Then, denoting the center of B_1 by c,

$$R = |pq| \leq |pc| + |cq| \leq R/2 + R/2 = R$$

and, therefore, $|cp| = |cq| = R/2$, which contradicts Lemma 1.

Thus, the diameter of S_1 is less than R. By the same argument, the diameter of S_2 is less than R. If we denote the size of S_1 by m, then it follows that $DP(S)$ is at most $m(n-m)$. Since the function $m(n-m)$, for $0 \leq m \leq n$, is maximized when $m = n/2$, the proof is complete. ∎

The proofs of the upper bounds on the weights of $MST(S)$ and $ANN(S)$ will be by induction on the size of the metric space S. We have to be careful, however, because the doubling dimension d' of a subset S' of S may be larger than the doubling d dimension of S. The reason is that the value of d' is determined by only considering balls that are centered at points of S'.

To give an example, let $n = m^2$ and let S' be a metric space of size n, in which the distance between any two distinct points is equal to 2. The doubling dimension d' of S' is equal to $d' = \log n = 2 \log m$. Partition S' into subsets $S'_1, S'_2, \ldots, S'_m$, each consisting of m points. Define $S = S' \cup \{p_1, p_2, \ldots, p_m\}$, where the p_i's are new points. For each i with $1 \leq i \leq m$, and for each q in S'_i, we define $|p_i q| = |q p_i| = 1$. For any other pair of distinct points in S, their distance is defined to be 2. Observe that S is a metric space.

We claim that the doubling dimension d of S is equal to $d = \log(m+1)$. To prove this claim, let $\rho > 0$ be a real number and let B be a ball in S of radius ρ. If $\rho \geq 2$, then B is covered by m balls of radius $\rho/2$, centered at the points $p_1, p_2, \ldots, p_m$. If $1 \leq \rho < 2$, then B contains at most $m+1$ points and, therefore, we can cover B by $m+1$ balls of radius $\rho/2$. Finally, if $\rho < 1$, then B is a singleton set and we can cover B by one ball of radius $\rho/2$. Thus, d is indeed equal to $\log(m+1)$.

For large values of m, the ratio d'/d converges (from below) to 2. The next lemma states that this ratio can never be larger than 2.

Lemma 2 *Let S be a finite metric space of doubling dimension d and let S' be a non-empty subset of S. Then the doubling dimension of the metric space S' is at most $2d$.*

Proof. Let c be a point of S', let $\rho > 0$ be a real number, and let $B' = \{p \in S' : |cp| \leq \rho\}$. Thus, B' is a ball in the metric space S'. Consider the corresponding ball B in S, i.e., $B = \{p \in S : |cp| \leq \rho\}$. Since S has doubling dimension d, we can cover B by balls $B_1, B_2, \ldots, B_k$, all having radius $\rho/4$, where $k \leq 2^{2d}$. For $1 \leq i \leq k$, let c_i be the center of B_i.

If $c_i \in S'$, then we define B'_i to be the ball in S' of radius $\rho/2$ that is centered at c_i. If $c_i \notin S'$ and $B_i \cap S' \neq \emptyset$, then we define B'_i to be the ball in S' of radius $\rho/2$ that is centered at an arbitrary point of $B_i \cap S'$.

We claim that the collection of balls $\{B'_i : 1 \leq i \leq k, B_i \cap S' \neq \emptyset\}$ cover the ball B'. To prove this, let p be a point in B'. Then $p \in B$ and, thus, there is an index i such that $p \in B_i$. If $c_i \in S'$, then $p \in B'_i$. Otherwise, let c'_i be the center of B'_i. Since

$$|c'_i p| \leq |c'_i c_i| + c_i p| \leq \rho/4 + \rho/4 = \rho/2,$$

it follows that $p \in B'_i$.

Thus, we have covered the ball B' in S' by at most 2^{2d} balls (again, in S') of radius $\rho/2$. ∎

Theorem 2 *Let S be a metric space consisting of n points, let d be the doubling dimension of S, and let R be the radius of S.*

1. *If $d = 1$, then $wt(MST(S)) \leq 2R \log n$.*
2. *If $d > 1$, then $wt(MST(S)) \leq 12Rn^{1-1/d}$.*

Proof. We start by considering the case when $d = 1$. We will prove the following claim by induction on the size of V: Let V be a non-empty subset of S and let R_V be the radius of V. Then

$$wt(MST(V)) \leq 2R_V \log |V|. \tag{1}$$

By taking $V = S$, this claim will prove the first part of the theorem.

If V is a singleton set, then (1) obviously holds. Assume that V contains at least two elements and, further, assume that (1) holds for all non-empty subsets of S having less than $|V|$ elements. Let B be a ball of radius R_V that contains V. Since the doubling dimension of S is equal to one, we can cover B by two balls B_1 and B_2, both having radius $R_V/2$. Note that the centers of B_1 and B_2 are not necessarily points of V. Define $V_1 = V \cap B_1$ and $V_2 = V \setminus V_1$. Then both V_1 and V_2 are non-empty (otherwise, the radius of V would be less than R_V). Let T be the spanning tree of V consisting of $MST(V_1)$, $MST(V_2)$, and one edge joining an arbitrary point of V_1 with an arbitrary point of V_2. Let R_{V_1} and R_{V_2} be the radius of V_1 and V_2, respectively. Then $R_{V_1} \leq R_V/2$ and $R_{V_2} \leq R_V/2$. Since the diameter of V is at most $2R_V$, and by applying the induction hypothesis, we obtain

$$\begin{aligned} wt(T) &\leq 2R_{V_1} \log |V_1| + 2R_{V_2} \log |V_2| + 2R_V \\ &\leq R_V \left(\log |V_1| + \log |V_2| + 2 \right). \end{aligned}$$

Since the function $-\log x$ is convex for $x > 0$, and since $|V_1| + |V_2| = |V|$, it follows that

$$wt(T) \leq R_V \left(\log(|V|/2) + \log(|V|/2) + 2 \right) = 2R_V \log |V|.$$

Since $wt(MST(V)) \leq wt(T)$, we have shown that (1) holds for V.

In the rest of the proof, we consider the case when $d > 1$. We will prove the following claim, again by induction on the size of V: Let V be a non-empty subset of S and let R_V be the radius of V. Then

$$wt(MST(V)) \leq 12R_V|V|^{1-1/d} - 12R_V. \tag{2}$$

By taking $V = S$, this claim will prove the second part of the theorem.

If V is a singleton set, then $wt(MST(V)) = R_V = 0$ and (2) holds. Assume that V consists of two points p and q. Then $wt(MST(V)) = |pq| \leq 2R_V$. Thus, (2) holds if we can show that $2 \leq 12 \cdot 2^{1-1/d} - 12$. This inequality can be rewritten as $7 \cdot 2^{1/d} \leq 12$. The latter inequality holds, because $d \geq \log 3$.

From now on, we assume that V contains at least three elements and (2) holds for all non-empty subsets of S having less than $|V|$ elements. Let B be a ball of radius R_V that contains V. Since the doubling dimension of S is equal to d, we can cover $B \cap V$ by balls $B_1, B_2, \ldots, B_k$, all having radius $R_V/2$, where $k \leq 2^d$. Observe that, since the radius of V is equal to R_V, we have $k \geq 2$. For $1 \leq i \leq k$, let V_i be the set of points of V that are in the ball B_i. In case a point of V is contained in more than one ball, we put it in exactly one subset V_i. Thus, the sets $V_1, V_2, \ldots, V_k$ are pairwise disjoint and their union is equal to V. We may also assume that none of the sets V_i is empty. We next claim that we may assume that $k \geq 3$. Indeed, if $k = 2$, then, since $|V| \geq 3$, one of the sets V_1 and V_2, say V_2, contains at least two points. We partition V_2 into two disjoint non-empty subsets, which we call V_2 and V_3, and, thus, we can set $k = 3$.

Let i be an index with $1 \leq i \leq k$, and let R_{V_i} be the radius of the set V_i. Then $R_{V_i} \leq R_V/2$. By the induction hypothesis, we have

$$wt(MST(V_i)) \leq 12 R_{V_i} \left(|V_i|^{1-1/d} - 1 \right).$$

Since $|V_i|^{1-1/d} - 1 \geq 0$, it follows that

$$wt(MST(V_i)) \leq 6 R_V \left(|V_i|^{1-1/d} - 1 \right).$$

For each i with $1 \leq i \leq k$, choose an arbitrary point in V_i, and let T' be an arbitrary spanning tree of these k points. Since the diameter of V is at most $2R_V$, we have

$$wt(T') \leq 2(k-1)R_V \leq 2kR_V.$$

Let T be the spanning tree of V consisting of T' and $MST(V_1)$, $MST(V_2)$, ..., $MST(V_k)$. Then

$$\begin{aligned} wt(T) &\leq \sum_{i=1}^{k} 6R_V \left(|V_i|^{1-1/d} - 1 \right) + 2kR_V \\ &= 6R_V \sum_{i=1}^{k} |V_i|^{1-1/d} - 4kR_V \\ &\leq 6R_V \sum_{i=1}^{k} |V_i|^{1-1/d} - 12R_V. \end{aligned}$$

Since the function $-x^{1-1/d}$ is convex for $x > 0$, and since $\sum_{i=1}^{k} |V_i| = |V|$, it follows that

$$\sum_{i=1}^{k} |V_i|^{1-1/d} \leq \sum_{i=1}^{k} \left(|V|/k \right)^{1-1/d} = k^{1/d} |V|^{1-1/d} \leq 2|V|^{1-1/d}.$$

Thus, we have

$$wt(T) \leq 12R_V|V|^{1-1/d} - 12R_V.$$

Since $wt(MST(V)) \leq wt(T)$, we have shown that (2) holds for V. ■

Theorem 3 *Let S be a metric space consisting of n points, where $n \geq 2$, let d be the doubling dimension of S, and let R be the radius of S.*

1. *If $d = 1$, then $wt(ANN(S)) \leq 4R$.*
2. *If $d > 1$, then $wt(ANN(S)) \leq 12Rn^{1-1/d}$.*

Proof. Since $ANN(S)$ is a subgraph of $MST(S)$, the second claim follows from Theorem 2. Assume that $d = 1$. The first claim will follow from the following more general claim: For any subset V of S with $|V| \geq 2$ and having radius R_V,

$$\sum_{p \in V} |p, NN_S(p)| \leq 4R_V. \tag{3}$$

The proof is by induction on the size of V. If V consists of two points a and b, then $\sum_{p \in V} |p, NN_S(p)| \leq 2|ab| \leq 4R_V$, thus (3) holds.

Assume that V contains at least three elements and, further, assume that (3) holds for all subsets of S having at least two and less than $|V|$ elements. Let B be a ball of radius R_V that contains V. We cover B by two balls B_1 and B_2, both having radius $R_V/2$, and define $V_1 = V \cap B_1$ and $V_2 = V \setminus V_1$. Observe that both V_1 and V_2 are non-empty. Let R_{V_1} and R_{V_2} be the radius of V_1 and V_2, respectively. Then $R_{V_1} \leq R_V/2$ and $R_{V_2} \leq R_V/2$.

If both V_1 and V_2 contain at least two points, then, using the induction hypothesis,

$$\begin{aligned}\sum_{p \in V} |p, NN_S(p)| &= \sum_{p \in V_1} |p, NN_S(p)| + \sum_{p \in V_2} |p, NN_S(p)| \\ &\leq 4R_{V_1} + 4R_{V_2} \\ &\leq 4R_V.\end{aligned}$$

Assume that one of V_1 and V_2, say V_1, consists of one point a. Since the set V has radius R_V, we have $|a, NN_S(a)| \leq R_V$. Combining this with the induction hypothesis, we obtain

$$\sum_{p \in V} |p, NN_S(p)| = |a, NN_S(a)| + \sum_{p \in V_2} |p, NN_S(p)| \leq R_V + 4R_{V_2} \leq 4R_V.$$

This completes the proof. ∎

3 The lower bounds

In this section, we show that the upper bounds in Theorems 1, 2 and 3 are tight. In fact, we give an example of one class of metric spaces that achieve the upper bounds in all these theorems. This class is obtained by repeatedly applying the following transformation:

Let $m \geq 2$ be an integer and assume we are given a metric space S' consisting of m points. Let d be the doubling dimension of S', let R' be the radius of S', and assume that the diameter of S' is equal to R'.

We define a new metric space S in the following way. For each i with $1 \leq i \leq 2^d$, let S'_i be a copy of the metric space S', and let S be the union of these copies. For any two points p and q in S, if they are in the same copy S'_i, then the distance $|pq|$ (in S) is defined to be their distance in S'_i. Otherwise, we define $|pq|$ to be $2R'$.

Let n be the size of S, so that $n = 2^d m$. The following claims can easily be verified. The triangle inequality holds for the set S and, thus, S is a metric space. Both the diameter and the radius of S are equal to $R = 2R'$ and

$$DP(S) = \binom{2^d}{2} m^2 = \frac{1}{2}\left(1 - \frac{1}{2^d}\right) n^2. \tag{4}$$

Since the diameter of each set S'_i is equal to R', and since distances between points in S'_i and points in S'_j (for $j \neq i$) are equal to $2R'$, we have

$$wt(ANN(S)) = 2^d \cdot wt(ANN(S')). \tag{5}$$

By running Kruskal's minimum spanning tree algorithm on S, it follows that

$$wt(MST(S)) = (2^d - 1)R + 2^d \cdot wt(MST(S')). \tag{6}$$

We finally claim that the doubling dimension of S is equal to d. To prove this, let c be a point in S, let $\rho > 0$ be a real number, and consider the ball $B = \{p \in S : |cp| \leq \rho\}$.

First assume that $\rho < 2R'$. Then B is completely contained in one of the sets S'_i. Since the doubling dimension of S'_i is equal to d, and since distances in S'_i are the same as distances in S, it follows that B can be covered by at most 2^d balls of radius $\rho/2$.

Now assume that $\rho \geq 2R'$. Since the radius of S'_i is equal to R', there exists a ball B'_i of radius R' that contains S'_i. The union of these balls B'_i, for $1 \leq i \leq 2^d$, cover the entire set S and, thus, it covers the ball B. Clearly, each such ball has radius at most $\rho/2$. Thus, also in this case, B can be covered by at most 2^d balls of radius $\rho/2$.

Theorem 4 *Let $\lambda \geq 2$ be an integer, let $d = \log \lambda$, let k be a positive integer, and let $n = 2^{kd}$. There exists a metric space S of size n and doubling dimension d, such that the following are true:*

1. *Both the radius and diameter of S are equal to $R = 2^k$.*
2. *$DP(S) = \frac{1}{2}\left(1 - 1/2^d\right) n^2$. In particular, if $d = 1$, then $DP(S) = n^2/4$.*
3. *$wt(ANN(S)) = 2Rn^{1-1/d}$. In particular, if $d = 1$, then $wt(ANN(S)) = 2R$.*
4. *If $d = 1$, then $wt(MST(S)) = R \log n$.*
5. *If $d > 1$, then $wt(MST(S)) \geq 2Rn^{1-1/d}$.*

Proof. Let S' be the metric space of size $m = 2^d$, in which the distance between any two distinct points is equal to 2. The doubling dimension of S' is equal to d, and both its radius and diameter are equal to $R' = 2$.

By applying the above construction $k-1$ times, we obtain a metric space S of size $n = 2^{kd}$, doubling dimension d, radius $R = 2^k$ and diameter $R = 2^k$, proving the first claim. The second claim follows from (4).

Since $wt(ANN(S')) = 2m = 2 \cdot 2^d$, it follows from (5) that $wt(ANN(S)) = 2 \cdot 2^{kd} = 2n$. Since $n = 2^{kd}$ and $R = 2^k$, we have $Rn^{1-1/d} = n$. Therefore, we have $wt(ANN(S)) = 2Rn^{1-1/d}$, proving the third claim.

The fourth claim follows from the recurrence in (6). Finally, the fifth claim follows from the third one, because $MST(S)$ contains $ANN(S)$. ∎

4 Concluding remarks

We have shown that metric spaces of low doubling dimension are very different than low-dimensional Euclidean spaces. In this final section, we present some more consequences of Lemma 1 for metric spaces of doubling dimension 1.

Lemma 3 *Let S be a finite metric space of doubling dimension* 1*, and let p, q, and r be three points in S such that (p, q) is an edge in $MST(S)$ and $|pq| \leq |pr| \leq |qr|$. Then $|pr| \geq 2|pq|$.*

Proof. Let $\ell = |pq|$, $\ell' = |pr|$, and $\ell'' = |qr|$. Thus, we have $\ell \leq \ell' \leq \ell''$.

Let B be the ball with center p and radius ℓ'. This ball contains the points p, q, and r. Cover B by balls B_1 and B_2 centered at c_1 and c_2, respectively, and having radius $\ell'/2$. We may assume without loss of generality that p is contained in B_1.

Assume that r is in B_1. Then

$$\ell' = |pr| \leq |pc_1| + |c_1 r| \leq \ell'/2 + \ell'/2 = \ell'$$

and, therefore, $|c_1 p| = |c_1 r| = \ell'/2$ and $|pr| = \ell'$, contradicting Lemma 1. Thus, r is in B_2.

If we assume that q is in B_2, then, by a similar argument, it follows that $|c_2 q| = |c_2 r| = \ell''/2$ and $|qr| = \ell''$, contradicting Lemma 1. Thus, q is in B_1.

We have shown that both p and q are in B_1, and that r is in B_2. Let $T = MST(S)$. By removing the edge (p, q) from T, we obtain two trees T_1 and T_2, where p is a node in T_1 and q is a node in T_2. We may assume without loss of generality that c_1 is a node in T_1. Observe that (c_1, q) is not an edge in T.

Let T' be the tree obtained from T by replacing the edge (p, q) by the edge (c_1, q). Since T is a minimum spanning tree of S, it follows that $\ell = |pq| \leq |c_1 q|$. On the other hand, since q is in B_1, we have $|c_1 q| \leq \ell'/2$. It follows that $\ell' \geq 2\ell$. ■

We now use Lemma 3 to prove the following claims. First, by following any (directed) path in $ANN(S)$, edges decrease in weight by a factor of at least 2. Second, for any point in S, any two incoming edges in $ANN(S)$ differ in weight by a factor of at least 2. Finally, any two edges in $MST(S)$ that share a point differ in length by a factor of at least 2.

Lemma 4 *Let S be a finite metric space of doubling dimension* 1*, and let p, q, and r be three pairwise distinct points in S.*

1. *If $q = NN_S(p)$ and $p = NN_S(r)$, then $|pr| \geq 2|pq|$.*

2. *If* $p = NN_S(q)$*,* $p = NN_S(r)$*, and* $|pq| \leq |pr|$*, then* $|pr| \geq 2|pq|$*.*

3. *If both* (p, q) *and* (p, r) *are edges in* $MST(S)$ *and* $|pq| \leq |pr|$*, then* $|pr| \geq 2|pq|$*.*

Proof. The assumptions in the first and second claims imply that (p, q) is an edge in $MST(S)$ and $|pq| \leq |pr| \leq |qr|$. Therefore, these two claims follow from Lemma 3.

For the third claim, observe that (q, r) is not an edge in $MST(S)$. It follows that $|pq| \leq |pr| \leq |qr|$ and, therefore, the claim also follows from Lemma 3. ∎

References

[1] P. Assouad. Plongements lipschitziens dans $\mathbb{R}^N$. *Bulletin de la Société Mathématique de France*, 111:429–448, 1983.

[2] L. Few. The shortest path and the shortest road through n points in a region. *Mathematika*, 2:141–144, 1955.

[3] S. Har-Peled and M. Mendel. Fast construction of nets in low-dimensional metrics and their applications. *SIAM Journal on Computing*, 35:1148–1184, 2006.

[4] J. Heinonen. *Lectures on Analysis on Metric Spaces.* Springer-Verlag, Berlin, 2001.

[5] J. Pach and P. K. Agarwal. *Combinatorial Geometry.* John Wiley & Sons, New York, 1995.

[6] M. Smid. The weak gap property in metric spaces of bounded doubling dimension. In *Efficient Algorithms, Essays Dedicated to Kurt Mehlhorn on the Occasion of His 60th Birthday*, volume 5760 of *Lecture Notes in Computer Science*, pages 275–289, Berlin, 2009. Springer-Verlag.

[7] W. D. Smith. *Studies in Computational Geometry Motivated by Mesh Generation.* Ph.D. thesis, Priceton University, Princeton, NJ, 1988.

[8] J. M. Steele and T. L. Snyder. Worst-case growth rates of some classical problems of combinatorial optimization. *SIAM Journal on Computing*, 18:278–287, 1989.