On the Impossibility of Dimension Reduction for Doubling Subsets of $\ell_p$

Yair Bartal*  
Hebrew University  
yair@cs.huji.ac.il

Lee-Ad Gottlieb  
Ariel University  
leead@ariel.ac.il

Ofer Neiman†  
Ben-Gurion University of the Negev  
neimano@cs.bgu.ac.il

ABSTRACT
A major open problem in the field of metric embedding is the existence of dimension reduction for $n$-point subsets of Euclidean space, such that both distortion and dimension depend only on the doubling constant of the pointset, and not on its cardinality. In this paper, we negate this possibility for $\ell_p$ spaces with $p > 2$. In particular, we introduce an $n$-point subset of $\ell_p$ with doubling constant $O(1)$, and demonstrate that any embedding of the set into $\ell_d^2$ with distortion $D$ must have $D \geq \Omega \left( \left( \frac{\log n}{\epsilon} \right)^{\frac{1}{2} - \frac{1}{p}} \right)$.

Categories and Subject Descriptors
F.2.2 [Nonnumerical Algorithms and Problems]: Geometrical problems and computations

General Terms
Algorithms, Theory

Keywords
Dimension reduction, Doubling metrics, Embedding

1. INTRODUCTION
Dimension reduction is one of the fundamental tools in algorithms design and a host of related fields. A particularly celebrated result in this area is the Johnson-Lindenstrauss Lemma [31], which demonstrates that any $n$-point subset of $\ell_2$ can be embedded with arbitrarily small distortion $1 + \epsilon$ into $\ell_d^2$ with $d = O(\log n/\epsilon^2)$. The JL-Lemma has found applications in such varied fields as machine learning [6, 8], compressive sensing [9], nearest-neighbor (NN) search [36], information retrieval [54] and many more.

A limitation of the JL-Lemma is that it is quite specific to $\ell_2$, and in fact there are lower bounds that rule out dimension reduction for the spaces $\ell_1$ and $\ell_\infty$. Yet essentially no non-trivial bounds are known for $\ell_p$ when $p \not\in \{1, 2, \infty\}$. The prospect of dimension reduction for $\ell_p$ would imply efficient algorithms for NN search and related proximity problems such as clustering, distance oracles and spanners.

The doubling constant of a metric space $(X,d)$ is the minimal $\lambda$ such that any ball of radius $2r$ can be covered by $\lambda$ balls of radius $r$, and the doubling dimension of $(X,d)$ is defined as $\log_2 \lambda$. A family of metrics is called doubling if the doubling constant of each of its members is bounded by some constant. The doubling dimension is a measure of the intrinsic dimensionality of a point set. In the past decade, it has been used in the development and analysis of algorithms for fundamental problems such as nearest neighbor search [34, 13, 19] and clustering [3, 21], for graph problems such as spanner construction [22, 17, 20, 26], the traveling salesman problem [53, 11], and routing [32, 52, 2, 33], and in machine learning [15, 23]. Importantly, it has also been observed that the doubling dimension often bounds the quality of embeddings for a point set, in terms of distortion and dimension [7, 27, 1, 18, 12, 24].

It is known that the dimension bounds of the Johnson-Lindenstrauss Lemma are close to optimal. A simple volume argument suggests that the set of $n$ standard unit vectors in $\mathbb{R}^n$ requires dimension at least $\Omega(\log n)$ to embed into any Euclidean embedding with distortion $D$. Alon [4] extended this lower bound to the low distortion regime and demonstrated that any embedding with $1 + \epsilon$ distortion requires $\Omega(\log n/(\epsilon^2 \log(1/\epsilon)))$ dimensions, thus showing that the Johnson-Lindenstrauss Lemma is nearly tight. However, this set of vectors has very high intrinsic dimension, as the doubling constant is $n$. Hence, it is only natural to ask the following:

QUESTION 1. Do subsets of $\ell_2$ with constant doubling dimension embed into constant dimensional space with low distortion?

This open question was first raised by [39, 27], and is considered among the most important and challenging problems in the study of doubling spaces [1, 18, 24, 46, 47]. In this work, we demonstrate that the best
work, we consider the natural counterpart of Question 1 for \( \ell_p \) spaces \((p > 2)\), and resolve our question in the negative:

**Theorem 1.** For any \( p > 2 \) there is a constant \( c = c(p) \) such that for any positive integer \( n \), there is a subset \( A \subseteq \ell_p^n \) of cardinality \( n \) with doubling constant \( O(1) \), such that any embedding of \( A \) into \( \ell_p^n \) with distortion at most \( D \) satisfies

\[
D \geq \Omega \left( \left( \frac{c \log n}{d} \right)^{\frac{1}{p}} \right).
\]

This result is the first non-trivial lower bound on dimension reduction in \( \ell_p \) spaces where \( p \not\in \{1, 2, \infty \} \). Additionally, it rules out a class of efficient algorithms for NN-search and the problems discussed above.

Note that the bound of Theorem 1 suggests that for sub-logarithmic dimension, the distortion must be non-constant.

The first impossibility result on dimension reduction in \( \ell_p \) spaces was given by [1, 18], and asserts that any \( n \)-point doubling metric can be embedded into \( \ell_p^n \) (for \( \log n \leq d \leq \log n \)) with distortion \( D \) at most

\[
D \leq O \left( \frac{\log n \log((\log n)/d)}{d^{1-1/p}} \right).
\]

There is no known improved upper bound on the embedding even if the metric lies in (high dimensional) \( \ell_p \) space.

**Techniques.**

While there exist numerous techniques for obtaining dimension reduction lower bounds in \( \ell_1 \) (see related work below), these all seem to be very specific to \( \ell_1 \) and fail for \( p \notin \{1, \infty \} \). Instead, we present a combinatorial proof for Theorem 1, which utilizes a new method based on potential functions.

The subset \( A \) is based on a recursive graph construction which is very popular for obtaining distortion-dimension tradeoffs [37, 48, 14, 27, 42, 40, 41, 5, 43, 50]. In these constructions one starts with a small basic graph, then in each iteration replaces every edge with the basic graph. In most constructions, the basic graph is very simple (e.g. a 4-cycle induces the so-called diamond graph) and is often a series-parallel graph. This is very useful for \( \ell_1 \), as [28, 16] showed these graphs embed to \( \ell_1 \) with constant distortion. However, these recursive graphs often require distortion \( \Omega((\log n)/p) \) for embedding into \( \ell_p \) with \( p > 2 \) [27], where \( n \) is the number of vertices, so one cannot use them directly. The novelty in this work is that the instance we produce is not a graph, but a certain subset of \( \ell_p \), which is inspired by the Laakso graph [37], the basic graph for which is depicted in Figure 1.

1.1 Related Work

Lafforgue and Naor [38] have concurrently proved the same result as in Theorem 1 using analytic tools, with a construction based on the Heisenberg group.

There are several results on embedding metric spaces with low intrinsic dimension into low dimensional normed space with low distortion: Assouad [7] showed that the snowflakes quantitative result that can be hoped for is an embedding with \( 1 + \epsilon \) distortion using \( O_{\epsilon}(\log \lambda) \) dimensions.

of doubling metrics\(^2\) embed with constant distortion into constant dimensional Euclidean space. In particular, this is a positive answer to Question 1 for this special case.

It is well known that an \( n \)-point metric space may require distortion \( \Omega(n^{1/\alpha}) \) for any embedding into Euclidean space [44]. It was shown by [27] that any doubling metric embeds with distortion only \( O((\log n)/\lambda) \), and that this is best possible. Their result was generalized by [35] to distortion \( O((\log n)^{1/(\log \lambda)} \), which was shown to be tight by [30].

As for low dimensional embedding, [1] showed that any doubling metric may be embedded with distortion \( O((\log n)^{1/\alpha}) \) (for any fixed \( \theta > 0 \)) into Euclidean space of dimension proportional to \( \log \lambda \), its intrinsic dimension. A trade-off between distortion and dimension for embedding doubling metrics was shown by [1, 18] as mentioned above. For doubling subsets of \( \ell_2 \), [24, 12] showed an embedding into constant dimensional \( \ell_2 \) with \( 1 + \epsilon \) distortion for a snowflake of the subset. For \( \alpha \)-snowflakes of arbitrary doubling metrics, [46] showed an embedding to Euclidean space where the dimension is a constant independent of \( \alpha \) (while the distortion, due to a lower bound of [51, 27, 40], must depend on it).

**Lower bounds on dimension reduction.**

The first impossibility result on dimension reduction in \( \ell_1 \) is due to [14], who showed that there exists a \( n \)-point subset of \( \ell_1 \) that requires \( d \geq n^{1/(1/p)} \) for any \( D \)-distortion embedding to \( \ell_p^n \). Following their work, there have been many different proofs and extensions of this result, using various techniques. The original [14] argument was based on linear programming and duality, then [42] gave a geometric proof. For the \( 1 + \epsilon \) distortion regime, [5] used combinatorial techniques to show that the dimension must be at least \( n^{1/(1-\epsilon)} \) (and also gave a different proof of the original result). Recently [50] applied an information theoretic argument to reprove the results of [14, 5]. As for linear dimension reductions, [40] showed a strong lower bound for \( \ell_p \) with \( p \neq 2 \).

The instances used by the papers mentioned above are based on recursive graph constructions. The papers of [14, 42] used the diamond graph, which has high doubling constant, but [40] showed that their proof can be extended to the Laakso graph, yielding essentially the same result but for a subset of low doubling dimension. For the \( \ell_\infty \) space, there are also strong lower bounds,\(^3\) which are based on large girth graphs [45] measure concentration [49] and geometric arguments [40].

There are few positive results for \( p \neq 2 \), such as [36] who showed that \( \ell_1 \) admits dimension reduction when the aspect ratio of the point set is bounded, and [10] used \( p \)-stable distributions to obtain similar results for all \( 1 < p < 2 \), and the Mazur map to obtain a (relatively high-distortion bound) for \( p > 2 \). A weak form of dimensionality reduction in \( \ell_1 \) was shown by [29].

2. CONSTRUCTION

Our construction is based on the Laakso graph [37], but will lie in \( \ell_p \) space. Abusing notation, we will refer to a

\(^2\)For \( 0 < \alpha < 1 \), an \( \alpha \)-snowflake of a metric \((X,d)\) is the metric \((X, d^\alpha)\), that is, all distances are taken to power \( \alpha \).

\(^3\)For instance, the metric induced by an \( n \)-point expander graph (which is in \( \ell_\infty \) as any other finite metric), requires dimension at least \( n^{1/\alpha} \) in any \( D \) distortion embedding.
pair of points as an edge, where edges will have a level. Fix a parameter $0 < \epsilon < 1/10$, and for a positive integer $k$ we shall define recursively an instance $A_k = A_k(\epsilon) \subseteq \mathbb{R}^k$. Let $e_1, \ldots, e_k$ be the standard orthonormal basis. In each instance $A_i$ certain pairs of points will be the level $i$ edges. The initial instance $A_1$ consists of the two points $e_1$ and $-e_1$, which are a level 1 edge. $A_i$ is created from $A_{i-1}$ by adding to every level $i-1$ edge $\{a,b\}$, four new points $s,t,u,v$ as follows:

$$s = \frac{3a}{4} + \frac{b}{4},$$

$$t = \frac{a}{4} + \frac{3b}{4},$$

$$u = \frac{a}{2} + \frac{b}{2} + \epsilon \|a-b\|_p \cdot e_i,$$

$$v = \frac{a}{2} + \frac{b}{2} - \epsilon \|a-b\|_p \cdot e_i,$$

and we will have the following six level $i$ edges:

$\{a,s\}, \{s,u\}, \{s,v\}, \{u,t\}, \{v,t\}, \{t,b\}$. These edges are the child edges of $\{a,b\}$ (see Figure 1). We will refer to the pair $\{u,v\}$ as a diagonal. In the full version of the paper we show that our construction must fail for $p = 2$, because the instance $A_k(\epsilon)$ (for any value of $\epsilon > 0$) embeds into the plane with constant distortion if distances are measured by the Euclidean norm.

### 3. DISTORTION-DIMENSION TRADEOFF

Fix any positive integers $d, D, k$, a real $p > 2$, and let $\epsilon = \epsilon(d, D, p)$ be the parameter by which we construct the instance $A_k$. The precise value of $\epsilon$ will be determined later. Assume that there is a non-expansive embedding $f$ of $A_k$ into $\ell_p^2$ with distortion $D$, where for each $j \in [d]$ there is a map $f_j : A_k \to \mathbb{R}$ and $f = \bigoplus_{j=1}^d f_j$. We want to show a tradeoff between the distortion $D$ and the dimension $d$. The argument is based on the tension between the edges and diagonals - the diagonals tend to contract and the edges to expand. To make this intuition precise, we employ the following potential function for each edge $\{a,b\}$,

$$\Phi(a,b) = \frac{\|f(a) - f(b)\|_2^2}{\|a - b\|_p^2}. \quad (1)$$

Since the embedding is non-expansive, and the image has only $d$ coordinates, by the power mean inequality

$$\|a - b\|_p \geq \|f(a) - f(b)\|_p \geq \|f(a) - f(b)\|_2 \cdot d^{1/p-1/2}.$$ 

Raising to power 2 and rearranging we obtain that the potential of any edge $\{a,b\}$ is never larger than

$$\Phi(a,b) \leq d^{1-2/p} . \quad (2)$$

The main goal, which is captured in the following Lemma, is to show that for a suitable choice of $\epsilon$, the potential increases (additively) by at least some positive number $\alpha = (\epsilon/D)^2$ at every level. Using (2) it must be that

$$k \leq d^{1-2/p}/\alpha , \quad (3)$$

as otherwise the potential of some level $k$ edge will be at least $\alpha k > \alpha \cdot d^{1-2/p}/\alpha = d^{1-2/p}$. The intuition behind the proof of the Lemma is simple: if $\{a,b\}$ is a level $i-1$ edge with potential value $\phi$, then consider the diagonal $\{u,v\}$ created from it in level $i$. Since $u,v$ have the same distance to $a,b$, the maximum potential of a child edge of $\{a,b\}$ will be minimized if $u,v$ are embedded into the same point in space. But in order to provide sufficient contribution for the diagonal $\{u,v\}$ we must have $u,v$ spaced out, which then causes some edges to expand, and thus increases the potential. The technical part of the proof balances between the loss in the potential (incurred because our instance lies in $\ell_p$ space), and the gain to the potential arising from the fact that $\|f(u) - f(v)\|_p$ must be large enough.

**Lemma 1.** There exists a constant $c = c(p)$ depending only on $p$ such that when $\epsilon \leq d^{-1/p} \cdot D^{1-(p-2)/p} \cdot c$ the following holds. For any level $i-1$ edge $\{a,b\}$ with potential value $\phi = \Phi(a,b)$, there exists a level $i$ edge (a child of $\{a,b\}$) with potential value at least $\phi + (\epsilon/D)^2$.

**Proof.** Let $s,t,u,v$ be the four new points introduced in $A_i$ from the edge $\{a,b\}$ as described above. For each $j \in [d]$ and $u,v$ we shall define $\Delta_j(u), \Delta_j(v) \in \mathbb{R}$ to be the values such that

$$f_j(u) = \frac{f_j(a)}{2} + \frac{f_j(b)}{2} + \Delta_j(u) \cdot \|a - b\|_p \quad (4)$$

$$f_j(v) = \frac{f_j(a)}{2} + \frac{f_j(b)}{2} + \Delta_j(v) \cdot \|a - b\|_p . \quad (5)$$

In what follows we show that the sum of the squares of the values $\Delta_j(u), \Delta_j(v)$ must be large, because the diagonal $\{u,v\}$ has sufficient contribution. Since the embedding has distortion $D$ we have that $\|u - v\|_p / D \leq \|f(u) - f(v)\|_p$. As

![Figure 1: The basic instance for the Laakso graph](image-url)
\[\|u - v\|_p = 2\|a - b\|_p\] it must be that
\[2\epsilon/D \cdot \|a - b\|_p \leq \|f(u) - f(v)\|_p\]

\[\leq \|a - b\|_p \left(\sum_{j=1}^{d} |\Delta_j(u) - \Delta_j(v)|^p\right)^{1/p} \]

\[\leq \|a - b\|_p \left(\sum_{j=1}^{d} 2^{p-1} |\Delta_j(u)|^p + \sum_{j=1}^{d} 2^{p-1} |\Delta_j(v)|^p\right)^{1/p} ,
\]

where the last inequality holds as for \( p > 1 \), by the power mean inequality \(|x - y|^p \leq (|x| + |y|)^p \leq 2^{p-1}(|x|^p + |y|^p)\). Hence for at least one of \( u, v \), say w.l.o.g. for \( u \), it follows that
\[\left(\sum_{j=1}^{d} |\Delta_j(u)|^p\right)^{1/p} \geq \epsilon/D .\]

Using that the \( \ell_2 \) norm is larger than the \( \ell_p \) norm for \( p > 2 \) we get that
\[\sum_{j=1}^{d} |\Delta_j(u)|^2 \geq (\epsilon/D)^2 . \quad (6)
\]

Next we consider the following two quantities:
\[\Phi'(u, a) = \frac{4\|f(u) - f(a)\|_2^2}{\|a - b\|_2^p} , \quad \Phi'(u, b) = \frac{4\|f(u) - f(b)\|_2^2}{\|a - b\|_2^p} .\]

Note that by (4)
\[\Phi'(u, a) = \frac{4}{\|a - b\|_2^p} \sum_{j=1}^{d} \left( f_j(b) - f_j(a) \right)^2 + |\Delta_j(u) : \|a - b\|_p^2 \right) \]

\[= \Phi(a, b) + \sum_{j=1}^{d} 4|\Delta_j(u)|^2 + \sum_{j=1}^{d} 4(f_j(b) - f_j(a)) \cdot |\Delta_j(u)| \].

Similarly using (5)
\[\Phi'(u, b) = \frac{4}{\|a - b\|_2^p} \sum_{j=1}^{d} \left( f_j(b) - f_j(a) \right)^2 - |\Delta_j(u) : \|a - b\|_p^2 \right) \]

\[= \Phi(a, b) + \sum_{j=1}^{d} 4|\Delta_j(u)|^2 - \sum_{j=1}^{d} 4(f_j(b) - f_j(a)) \cdot |\Delta_j(u)| \].

As the two terms only differ by the sign before the term \( \sum_{j=1}^{d} 4(f_j(b) - f_j(a)) \cdot |\Delta_j(u)| \), we conclude that at least one of them, assume w.l.o.g. \( \Phi'(u, a) \), must be at least
\[\Phi'(u, a) \geq \Phi(a, b) + \sum_{j=1}^{d} 4|\Delta_j(u)|^2 . \quad (7)
\]

Now we shall consider
\[\Phi'(a, s) = \frac{16\|f(a) - f(s)\|_2^2}{\|a - b\|_2^p} , \quad \Phi'(u, s) = \frac{16\|f(u) - f(s)\|_2^2}{\|a - b\|_2^p} .\]

Observe that since \( \|a - s\|_p = \frac{1}{2}\|a - b\|_p \) we have that
\[\Phi'(a, s) = \Phi(a, s) . \quad (8)
\]

Since \( e, \) is a unit vector orthogonal to the subspace in which \( s \) lies, we have that \( \|u - s\|_p = \frac{1}{2}\|a - b\|_p (1 + (4\epsilon)^p)^{1/p} \), and then
\[\Phi'(u, s) = (1 + (4\epsilon)^p)^{1/p} \cdot \Phi(u, s) . \quad (9)
\]

Next, note that
\[\Phi'(u, a) = \frac{4\|f(u) - f(s) + f(s) - f(a)\|_2^2}{\|a - b\|_2^p} \]

\[\geq \frac{8 \|f(u) - f(s)\|_2^2 + \|f(s) - f(a)\|_2^2}{\|a - b\|_2^p} \]

\[\geq \frac{1}{2} \Phi(u, s) + \Phi'(a, s) ,
\]

so one of \( \Phi'(u, s), \Phi'(a, s) \) is at least as large as \( \Phi'(u, a) \). If it is the case that \( \Phi'(a, s) \geq \Phi'(u, a) \), then the assertion of the Lemma is proved for the level \( l \) edge \( \{a, s\} \), because by (8), (6) and (7)
\[\Phi(a, s) = \Phi'(a, s) \geq \Phi(a, u) \geq \Phi(a, b) + 4\epsilon/D^2 ,
\]

and by (9)
\[\Phi(u, s) \geq \Phi(a, b) + 4\epsilon/D^2 \]

\[\geq 1 + (4\epsilon)^p/2 \]

\[\Phi(u, s) \geq \Phi(a, b) + 4\epsilon/D^2 . \]

We will use that for \( 0 < x < 1/2, \epsilon^x \leq 1 + 2x, \) then as \( \epsilon < 1/10, \)
\[1 + (4\epsilon)^p/2 \leq \epsilon^{2(4\epsilon)^p/2} \leq 1 + 4(4\epsilon)^p/2 \leq 1 + (6\epsilon)^p \leq 1 / (1 - (6\epsilon)^p) .\]

Using this, that \( 1 - (6\epsilon)^p > 1/2 \) and by (2),
\[\Phi(u, s) \geq \Phi(a, b) + 4\epsilon/D^2 \cdot (1 - (6\epsilon)^p) \]

\[\geq \Phi(a, b) - (6\epsilon)^p \cdot d^{(1-2/p)} + 2\epsilon/D^2 .\]

Recall that \( \epsilon \leq d^{-1/p} \cdot D^{-2/(p-2)}/c \), and we can set \( c = \Phi(h^p/(p-2)) \) so that
\[2\epsilon/D^2 - (6\epsilon)^p \cdot d^{(1-2/p)} \geq \epsilon/D^2 ,
\]

which satisfies the assertion of the Lemma for the edge \( \{u, s\} \).

Finally, let us prove the main Theorem.

**Proof of Theorem 1.** The set \( A_k \) has a doubling constant \( O(1) \) as established in Section 4. The number of points in \( A_k \) is \( n = \Theta(6^k) \), so \( k = \Theta(\log n) \). Using Lemma 1 and (3) we have that
\[\Omega(\log n) \leq d^{1-2/p} \cdot (D/\epsilon)^2 \]

\[= d^{1-2/p} \cdot D^2 \cdot d^{2/p} \cdot D^4/(p-2) \cdot c^2 \]

\[\leq c^2 d \cdot D^{2p/(p-2)} .
\]
So if one fixes the dimension $d$, we conclude that the distortion must be at least

$$D \geq \Omega_p \left( \frac{\log n}{d} \right)^{1/2 - 1/p}.$$  \hspace{2cm} (10)

\[\square\]

**Remark 1.** A similar calculation shows that if the embedding of $A_k \subseteq \ell_p^n$ is done into $\ell_p^n$ for some $q \geq 1$, then the distortion is at least

$$D \geq \Omega_p \left( \frac{\log n}{d} \right)^{1/2 - 1/p}.$$  \hspace{2cm} (11)

\section{A \textit{K} IS DOUBLING}

In this section we prove that $A_k$ is doubling. We first construct a graph $B_i$ very similar to $A_k$, with infinite number of points. The initial graph $B_0$ consists of a single line segment connecting two points $a, b$ of distance 1. Graph $B_1$ is the Laakso graph decomposition of the line, where child edges \{a, s\}, \{t, b\} have length $\frac{1}{2}$ and \{s, u\}, \{s, v\}, \{u, t\}, \{v, t\} have length $q = \frac{1 + (4^{1/p} + 1)}{2^{1/p}}$. (Note that since $\frac{1}{2} < \frac{1}{2^{1/p}} \leq q < \frac{1}{2^{1/p}} < \frac{1}{2^{1/p}}$.) Likewise, graph $B_j$ is formed from $B_{j-1}$ by applying the Laakso decomposition to edges with length in the range $[q^i, q^{i-1}]$. The new coordinate used by the Laakso decomposition on the edge is the same as the one used by the instance $A_k$ on the same edge; $B_i$ differs only in the order in which edges are decomposed.

It follows that all edges in $B_i$ have length in the range $\left[\frac{q^i}{\sqrt{2}}, q^{i-1} \right)$. Further, since $\epsilon < 1/10$, the distance from an edge in $B_i$ of length $r \in \left[\frac{q^i}{\sqrt{2}}, q^{i-1} \right)$ to all descendant points of that edge (in graphs $B_j, j > i$) is at most

$$C r \sum_{j=0}^{\infty} q^j \leq \frac{40}{29} r \leq \frac{3}{2} C r,$$

(11)

where the distance between two segments is the distance between the closest pair of points on the segments. We say that segments are disjoint if their intersection is empty. We have the following lemma:

**Lemma 2.** Let $e \in B_i$ be of length $r \in \left[\frac{q^i}{\sqrt{2}}, q^{i-1} \right)$. Then the distance from $e$ to any disjoint edge in $B_i$ is at least $\frac{\sqrt{2}}{3} r$.

**Proof.** We will prove part of the lemma by induction, with the base case being instance $B_1$. The distance between disjoint edges \{a, s\}, \{t, b\} of length $r = \frac{1}{2}$ is $\frac{1}{\sqrt{2}} r$. The distance between edges \{s, u\}, \{s, v\}, \{u, t\} (or edges \{s, v\}, \{u, t\}) of length $r = q$ is greater than half the distance from $u$ to $v$: $\epsilon = \frac{1}{\sqrt{2}} r > \frac{1}{\sqrt{2}} r$.

For the induction step, take edge $e \in B_i$ of length $r$ and consider its distance from $e' \in B_i$. Consider the respective parent edges of $e, e'$ — call them $p(e), p(e') \in B_{i-1}$ — and assume without loss of generality that $p(e)$ is not shorter than $p(e')$. Then the length of $p(e)$ in $r' \in \left[\frac{1}{2}, 4r \right]$ and the distance from $p(e)$ (respectively, $p(e')$) to the farthest point in $e$ (resp. $e'$) is at most $r'$. If $p(e)$ and $p(e')$ are disjoint, then by assumption their distance is at least $\frac{1}{\sqrt{2}} r'$, and so the distance from $e$ to $e'$ is at least $\frac{1}{\sqrt{2}} r' - 2r' > \frac{1}{\sqrt{2}} r$.

The above induction is valid only for $e, e'$ with disjoint parents, and we require a separate for intersecting parents.

We first consider the case where the parents $p(e), p(e')$ meet in a joint such as formed by edges \{s, u\}, \{s, v\} of the Laakso graph. Note that in this case the lengths of $p(e), p(e')$ are always equal, and the children of $p(e), p(e')$ use new coordinates. We decompose $p(e)$ into edges \{a', s\}, \{s', v\}, \{s', v', \{u', s', \{v', t', \{t', v'\}. Let $e$ be one of the longer children of $p(e)$, and $r$ its length. Aside from child edge \{a', s\} (which intersects $p(e')$), the distance from all child edges to $p(e')$ is at least the distance from $s'$ to $p(e')$. This in turn is greater than the distance from $s'$ to the (hypothetical) edge \{s, t\}, which is exactly $\frac{1}{\sqrt{2}} r$. The children of $p(e')$ can only be farther away than this distance, and a similar argument bounds the distance from \{a', s\} to all children of $p(e')$.

It remains to consider the cases where $p(e)$ and $p(e')$ form a joint similar to that of \{a, s\}, \{s, u\} or \{s, v\}, \{u, t\}. In either case the distance from the children of $p(e)$ to $p(e')$ and its children is at least a quarter of the length of $p(e)$.

We can then prove the following theorem:

**Lemma 3.** Let $E$ be a collection of disjoint edges in $B_i$ with lengths in the range $\left[\frac{q^i}{\sqrt{2}}, q\right]$ and inter-edge distance at most $30q$. Then $|E| = O(1)$.

**Proof.** Consider any two edges $e_1, e_2 \in E$, and let $p(e_1), p(e_2) \in B_i$, be their respective ancestors in level $j = i - 5$. We will show that $p(e_1), p(e_2)$ must intersect (or in fact the same edge): Let us assume $p(e_1)$ is not shorter than $p(e_2)$, and has length $r \in \left[\frac{q^i}{\sqrt{2}}, q\right]$. By the assertion of the lemma together with (11), the distance between ancestral edges $p(e_1), p(e_2) \in B_i$ is at most $30q + 2^\frac{1}{2} 3er < 2^\frac{1}{2} q + 3er \leq 32er$. However, by Lemma 2, if $p(e_1), p(e_2)$ are disjoint their distance must be at least $\frac{1}{\sqrt{2}} r > 3.3er$. So the ancestral edges intersect.

The maximum number of mutually adjacent edges in $B_i$ is 3, so there are at most 3 ancestral edges in $B_i$ for all edges in $E$. Since each edge can beget 6 child edges, $|E| \leq 3 \cdot 6^j$. \[\square\]

We can now bound the doubling dimension of $A_k$: Consider the maximal set $S \subseteq A_k$ of points with inter-point distances in the range $[\gamma, 2\gamma]$ (for any $\gamma$); then the doubling constant of $A_k$ is at most $|S|^2$ [25]. Choose $i$ such that $\frac{q^i}{\sqrt{2}} \leq 2\gamma < q^i$, and consider graph $B_i$. Associate each point $x \in S$ with its ancestral edge in $B_i$, and by (11) the distance of $x$ from $s(x)$ is less than $\frac{3}{2} q^i \leq \frac{1}{\sqrt{2}}$, thus the projected points have inter-point distance in the range $\left[\frac{1}{2}, \frac{1}{\sqrt{2}}\right]$. This distance range implies that any single edge of $B_i$ can have at 6 points projected upon it. By Lemma 3, the number of edges in $B_i$ with inter-point distance at most $\frac{q^i}{\sqrt{2}} \leq 30q$ is bounded by a constant, so $|S| = O(1)$.

**5. REFERENCES**

[1] Ittai Abraham, Yair Bartal, and Ofer Neiman. Embedding metric spaces in their intrinsic dimension. In Proceedings of the 19th annual ACM-SIAM symposium on Discrete algorithms, SODA ’08, pages 363–372, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[2] Ittai Abraham, Cyril Gavoille, Andrew V. Goldberg, and Dahlia Malkhi. Routing in networks with low doubling dimension. In 24th International Conference on Distributed Computing Systems (ICDCS), IEEE Computer Society Press, July 2006.
[33] G. Konjevod, A. W. Richa, D. Xia, and H. Yu. Compact routing with slack in low doubling dimension. In 26th annual ACM symposium on Principles of distributed computing, pages 71–80. ACM, 2007.

[34] Robert Krauthgamer and James R. Lee. Navigating nets: simple algorithms for proximity search. In SODA ’04: Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms, pages 798–807, Philadelphia, PA, USA, 2004. Society for Industrial and Applied Mathematics.

[35] Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: a new embedding method for finite metrics. Geometric and Functional Analysis, 15(4):839–858, 2005.

[36] Eyal Kushilevitz, Rafail Ostrovsky, and Yuval Rabani. Efficient search for approximate nearest neighbor in high dimensional spaces. In Proceedings of the thirtieth annual ACM symposium on Theory of computing, STOC ’98, pages 614–623, New York, NY, USA, 1998. ACM.

[37] Tomi J. Laakso. Plane with $A_\infty$-weighted metric not bi-Lipschitz embeddable to $\mathbb{R}^n$. Bull. London Math. Soc., 34(6):667–676, 2002.

[38] Vincent Lafforgue and Assaf Naor. A doubling subset of $L_p$ for $p > 2$ that is inherently infinite dimensional, 2013. preprint.

[39] Urs Lang and Conrad Plaut. Bilipschitz embeddings of metric spaces into space form. Geometriae Dedicata, 87(1-3):285–307, 2001.

[40] James R. Lee, Manor Mendel, and Assaf Naor. Metric structures in $l_1$: dimension, snowflakes, and average distortion. Eur. J. Comb., 26(8):1180–1190, November 2005.

[41] James R. Lee and Mohammad Moharrami. Bilipschitz snowflakes and metrics of negative type. In Proceedings of the 42nd ACM symposium on Theory of computing, STOC ’10, pages 621–630, New York, NY, USA, 2010. ACM.

[42] James R. Lee and Assaf Naor. Embedding the diamond graph in $l_p$ and dimension reduction in $l_1$. pages 745–747, 2004.

[43] James R. Lee and Anastasios Sidiropoulos. Near-optimal distortion bounds for embedding doubling spaces into $l_1$. In Proceedings of the 43rd annual ACM symposium on Theory of computing, STOC ’11, pages 765–772, New York, NY, USA, 2011. ACM.

[44] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. Combinatorica, 15(2):215–245, 1995.

[45] Jiri Matousek. Lectures on discrete geometry. Springer-Verlag, New York, 2002.

[46] Assaf Naor and Ofer Neiman. Assouad’s theorem with dimension independent of the snowflaking. Revista Matematica Iberoamericana, 28(4):1–21, 2012.

[47] Ofer Neiman. Low dimensional embedding of doubling metrics. In Proceedings of the 11th international conference on Approximation and Online Algorithms, WAOA, 2013.

[48] Ilan Newman and Yuri Rabinovich. A lower bound on the distortion of embedding planar metrics into euclidean space. In SCG ’02: Proceedings of the 18th annual symposium on Computational geometry, pages 94–96. ACM Press, 2002.

[49] Yuri Rabinovich. On average distortion of embedding metrics into the line and into $l_1$. In Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, STOC ’03, pages 456–462, New York, NY, USA, 2003. ACM.

[50] Oded Regev. Entropy-based bounds on dimension reduction in $l_1$. Israel Journal of Mathematics, pages 1–8, 2012.

[51] S. Semmes. On the nonexistence of bilipschitz parameterizations and geometric problems about $a_\infty$ weights. Revista Matematica Iberoamericana, 12:337–410, 1996.

[52] Aleksandrs Slivkins. Distance estimation and object location via rings of neighbors. In 24th Annual ACM Symposium on Principles of Distributed Computing (PODC), pages 41–50. ACM Press, July 2005. Appears earlier as Cornell CIS technical report TR2005-1977.

[53] Kunal Talwar. Bypassing the embedding: algorithms for low dimensional metrics. In STOC ’04: Proceedings of the thirty-sixth annual ACM symposium on Theory of computing, pages 281–290. ACM Press, 2004.

[54] S. S. Vempala. The Random Projection Method. American Mathematical Society, 2004.