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

Yair Bartal†  Lee-Ad Gottlieb‡  Ofer Neiman§

April 21, 2015

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 $\ell_p^d$ with distortion $D$ must have $D \geq \Omega\left(\left(\frac{\log n}{d}\right)^{\frac{2}{p}-\frac{1}{2}}\right)$.

*A preliminary version of this paper appeared in Proceedings of SOCG 2014
†Hebrew University. Email: yair@cs.huji.ac.il. Supported in part by a grant from the Israeli Science Foundation (1609/11).
‡Ariel University. Email: leead@ariel.ac.il.
§Ben-Gurion University of the Negev. Email: neimano@cs.bgu.ac.il. Supported in part by ISF grant No. (523/12) and by the European Union’s Seventh Framework Programme (FP7/2007-2013) under grant agreement n°303809.
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 [32], which demonstrates that any \( n \)-point subset of \( \ell_2 \) can be embedded with arbitrarily small distortion \( 1 + \epsilon \) into \( \ell_2^d \) 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 [37], information retrieval [55] 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 \( p \notin \{1, 2, \infty\} \). 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 [35, 13, 20] and clustering [3, 22], for graph problems such as spanner construction [23, 17, 21, 27], the traveling salesman problem [54, 11], and routing [33, 53, 2, 34], and in machine learning [15, 24]. 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, 28, 1, 18, 12, 25].

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 \( \Omega(\log_D n) \) 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 [40, 28], and is considered among the most important and challenging problems in the study of doubling spaces [1, 18, 25, 47, 48]. In this 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 \( A \subseteq \ell_p \) of cardinality \( n \) with doubling constant \( O(1) \), such that any embedding of \( A \) into \( \ell_p^d \) with distortion at most \( D \) satisfies

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

This result is the first non-trivial lower bound on dimension reduction in \( \ell_p \) spaces where \( p \notin \{1, 2, \infty\} \). 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. In the case \( p = \infty \), a result of [28] (see also [48]) states that any doubling metric embeds with

---

1 It follows from simple volume arguments that the best quantitative result that can be hoped for is an embedding with \( 1 + \epsilon \) distortion using \( O(\log \lambda) \) dimensions.
distortion $1 + \epsilon$ into $\ell_\infty^d$ with $d = \epsilon^{-O(1)} \log n$, so in this sense our result is tight. Additionally, a tradeoff between the distortion and dimension for embedding general doubling metrics into $\ell_p$ was given by [1, 18], and asserts that any $n$-point doubling metric can be embedded into $\ell_p^d$ (for $\log \log n \leq d \leq \log n$) 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 $\mathcal{A}$ is based on a recursive graph construction which is very popular for obtaining distortion-dimension tradeoffs [38, 49, 14, 28, 43, 41, 42, 5, 44, 51]. 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 [29, 16] showed these graphs embed to $\ell_1$ with constant distortion. However, these recursive graphs often require distortion $\Omega((\log^{1/p} n)$ for embedding into $\ell_p$ with $p > 2$ [28], 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 [38], the basic graph for which is depicted in Figure 1.

1.1 Related Work

Lafforgue and Naor [39] 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 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.

\(^2\)For $0 \leq \alpha < 1$, an $\alpha$-snowflake of a metric $(X, d)$ is the metric $(X, d^\alpha)$, that is, all distances are taken to power $\alpha$. 
It is well known that arbitrary $n$ point metrics may require distortion $\Omega(\log n)$ for any embedding into Euclidean space \cite{45}. It was shown by \cite{28} that any doubling metric embeds with distortion only $O(\sqrt{\log n})$, and that this is best possible. Their result was generalized by \cite{36} to distortion $O(\sqrt{\log n} \cdot \log \lambda)$, which was shown to be tight by \cite{31}. As for low dimensional embedding, \cite{1} showed that any doubling metric may be embedded with distortion $O(\sqrt{\log n})$ (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 \cite{1, 18} as mentioned above. For doubling subsets of $\ell_2$, \cite{25, 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, \cite{47} showed an embedding to Euclidean space where the dimension is a constant independent of $\alpha$ (while the distortion, due to a lower bound of \cite{52, 28, 41}, must depend on it).

**Lower bounds on dimension reduction**  The first impossibility result on dimension reduction in $\ell_1$ is due to \cite{14}, who showed that there exists an $n$-point subset of $\ell_1$ that requires $d \geq n^{\Omega(1/D^2)}$ for any $D$-distortion embedding to $\ell_1^d$. Following their work, there have been many different proofs and extensions of this result, using various techniques. The original \cite{14} argument was based on linear programming and duality, then \cite{43} gave a geometric proof. For the $1 + \epsilon$ distortion regime, \cite{5} used combinatorial techniques to show that the dimension must be at least $n^{1-O(1/(\log(1/\epsilon)))}$ (and also gave a different proof of the original result). Recently \cite{51} applied an information theoretic argument to reprove the results of \cite{14, 5}. As for linear dimension reductions, \cite{41} showed a strong lower bound for $\ell_p$ with $p \neq 2$, and \cite{19} showed a lower bound for $\ell_1$.

The instances used by the papers mentioned above are based on recursive graph constructions. The papers of \cite{14, 43} used the diamond graph, which has high doubling constant, but \cite{41} 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 \cite{46} measure concentration \cite{50} and geometric arguments \cite{41}.

There are few positive results for $p \neq 2$, such as \cite{37} who showed that $\ell_1$ admits dimension reduction when the aspect ratio of the point set is bounded, and \cite{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 \cite{30}.

**2 Construction**

Our construction is based on the Laakso graph \cite{38}, but will lie in $\ell_p$ space. Abusing notation, we will refer to a pair of points as an edge, where edges will have a level. Fix a parameter $0 < \epsilon < 1/17$, 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

\(^3\)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 $n^{1/O(D)}$ in any $D$ distortion embedding.
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 Section 5 we show that our construction must fail for $p = 2$, because the instance $A_{k}(\epsilon)$ (for sufficiently small value of $\epsilon > 0$) embeds into the plane with constant distortion.

### 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 $\ell_{p}^{d}$ with distortion $D$, where for each $j \in [d]$ there is a map $f_{j} : A_{k} \rightarrow \mathbb{R}$ 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 $\{a, b\}$,

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

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

$$\|a - b\|_{p} \geq \|f(a) - f(b)\|_{p} \geq \|f(a) - f(b)\|_{2} \cdot d^{1/p - 1/2}.$$\n
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 $\epsilon \leq d^{-1/p} \cdot D^{-2/(p-2)}/c$ the following holds. For any level $i - 1$ edge $\{a, b\}$ with potential value $\phi = \Phi(a, b)$, there exists a level $i$ edge (who is 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

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

In what follows we show that the sum of the squares of the values $\Delta_j(u), \Delta_j(v)$ must be large, because of the distortion requirement on the diagonal $\{u, v\}$. Since the embedding has distortion $D$ we have that $\|u - v\|_p/D \leq \|f(u) - f(v)\|_p$. As $\|u - v\|_p = 2\epsilon\|a - b\|_p$ it must be that

$$2\epsilon/D \cdot \|a - b\|_p \leq \|f(u) - f(v)\|_p \leq \|u - v\|_p/D \leq 2\epsilon\|a - b\|_p.$$  

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

$$\sum_{j=1}^{d} \Delta_j(u)^2 \geq (\epsilon/D)^2.$$  

Next we consider the following two quantities:

$$\Phi'(u, a) = \frac{4\|f(u) - f(a)\|_2^2}{\|a - b\|_p^2},$$
$$\Phi'(u, b) = \frac{4\|f(u) - f(b)\|_2^2}{\|a - b\|_p^2}.$$  

3
Note that by (4)
\[
\Phi'(u, a) = \frac{4}{\|a - b\|^2} \sum_{j=1}^{d} \left( \frac{f_j(b) - f_j(a)}{2} + \Delta_j(u) \cdot \|a - b\|_p \right)^2
\]
\[
= \Phi(a, b) + \sum_{j=1}^{d} 4\Delta_j(u)^2 + \sum_{j=1}^{d} \frac{4(f_j(b) - f_j(a)) \cdot \Delta_j(u)}{\|a - b\|_p}.
\]

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

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

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

Observe that since $\|a - s\|_p = \frac{1}{4} \|a - b\|_p$ we have that
\[
\Phi'(a, s) = \Phi(a, s). \tag{8}
\]

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

Next, note that
\[
\Phi'(u, a) = \frac{4\|f(u) - f(s) + f(s) - f(a)\|^2}{\|a - b\|^2}
\]
\[
\leq 8 \frac{\left(\|f(a) - f(s)\|^2 + \|f(s) - f(u)\|^2\right)}{\|a - b\|^2}
\]
\[
= \frac{1}{2} (\Phi'(u, s) + \Phi'(a, s)).
\]
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 $i$ edge $\{a, s\}$, because by (8), (6) and (7)

$$
\Phi(a, s) = \Phi'(a, s) \\
\geq \Phi'(u, a) \\
\geq \Phi(a, b) + \sum_{j=1}^{d} 4\Delta_j(u)^2 \\
\geq \Phi(a, b) + 4(\epsilon/D)^2.
$$

So from now on we focus on the case that $\Phi'(u, s) \geq \Phi'(u, a)$. The same calculation shows that

$$
\Phi'(u, s) \geq \Phi'(u, a) \geq \Phi(a, b) + 4(\epsilon/D)^2,
$$

and by (9)

$$
\Phi(u, s) \geq \frac{\Phi(a, b) + 4(\epsilon/D)^2}{(1 + (4\epsilon/p)^{2/p})}.
$$

We will use that for $0 < x < 1/2, e^x \leq 1 + 2x$, then as $\epsilon < 1/10$,

$$(1 + (4\epsilon/p)^{2/p}) \leq e^{2(4\epsilon/p)/p} \leq 1 + 4(4\epsilon/p)^{p/p} \leq 1 + (6\epsilon)^p \leq \frac{1}{1 - (6\epsilon)^p}.$$ 

Using this, the fact that $1 - (6\epsilon)^p > 1/2$, and by (2),

$$
\Phi(u, s) \geq \left( \Phi(a, b) + 4(\epsilon/D)^2 \right) \cdot (1 - (6\epsilon)^p) \\
\geq \Phi(a, b) - (6\epsilon)^p \cdot d^{1-2/p} + 2(\epsilon/D)^2.
$$

Recall that $\epsilon \leq (d^{-1/p} \cdot D^{-2/(p-2)})/c$, and we can set $c = 6p/(p-2)$ 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 $\{u, s\}$. \qed

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

$$
\Omega(\log n) \leq \frac{d^{1-2/p} \cdot (D/\epsilon)^2}{(d^{1-2/p} \cdot D^2 \cdot D^{4/(p-2)} \cdot c^2)} \\
= \frac{c^2 d \cdot D^{2p/(p-2)}}{(d^{1-2/p} \cdot D^2)}.
$$

So if one fixes the dimension $d$, we conclude that the distortion must be at least

$$
D \geq \Omega \left( \left( \frac{c' \log n}{d} \right)^{1/2 - 1/p} \right), \quad (10)
$$

where $c'$ is a constant that depends only on $p$. \qed

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

$$
D \geq \Omega \left( \frac{c' \log^{1/2 - 1/p} n}{d^{\max(q - 2, 2 - q)/2q}} \right),
$$

8
4 \( A_k \) is Doubling

In this section we prove that \( A_k \) is doubling. The basic reason is essentially the same as the reason for the Laakso graph; the edge \( \{a, s\} \) and \( \{t, b\} \) isolate the complexity of the recursive instance from the rest of the graph (see Figure 1). However, in our setting where the instance lies in \( \ell_p \), we must argue that the instance constructed recursively from some given segment \( \{a, b\} \), will be "sufficiently far" from the instances corresponding to other edges. The main difficulty arises since the diagonals \( u, v \) are only \( \epsilon \|a - b\|_p \) far from the segment of \( \{a, b\} \), so the instances created from the segments \( \{s, u\} \) and \( \{s, v\} \) are rather close, while we desire a doubling constant that is independent of \( \epsilon \).

We first construct a graph \( B_k \) 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 \( \frac{(1/4)^p + \epsilon^p)^{1/p}}{q} \). Note that since \( \epsilon < \frac{1}{17}, \frac{1}{4} < q < \frac{1}{17} + \frac{1}{17^2} < \frac{1.03}{4} \), where the minimum is achieved when \( p = \infty \) and the maximum as \( p \) approaches 2. Graph \( B_{i+1} \) is formed from \( B_i \) by applying the Laakso decomposition to edges with length in the range \( (q \cdot q^i, q^{i+1}) \), so that all edges in \( B_{i+1} \) have length in the range \( (4^{-1} \cdot q^i, q^{i+1}) \). More generally, all edges in \( B_j \) have length in the range \( (4^{-1} \cdot q^j, q^j) \), and among these only the longer edges of length \( (q \cdot q^j, q^{j+1}) \) will be decomposed in the next round. 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 – the new graph differs only in the order in which edges are decomposed.

By construction, the distance from edge \( e \in B_i \) of length \( r \in (\frac{q^j}{q}, q^j) \) to a child point of \( e \) in graph \( B_{i+1} \) is at most \( \epsilon r \). Then the distance from \( e \) to all descendant points in graphs \( B_j, j > i \), is at most

\[
\epsilon r \sum_{j=0}^{\infty} q^j \leq \left( 1 - \frac{1.03}{4} \right)^{-1} \epsilon r < 1.35 \epsilon r < \frac{3}{2} \epsilon r. \tag{11}
\]

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 edge \( e \in B_i \) be of length \( r \in (\frac{q^j}{q}, q^j) \), and set \( \epsilon < \frac{1}{17} \). Then the distance from \( e \) to any disjoint edge in \( B_i \) is at least \( \frac{\epsilon}{4q^{2j}} r \).

**Proof.** We will prove the lemma by induction, with the base case being instance \( B_1 \). Clearly, the closest edges in \( B_1 \) are pairs \( \{s, u\}, \{v, t\} \) and \( \{s, v\}, \{u, t\} \) of length \( r = q \). The distance between the edges is greater than \( \epsilon \) (half the distance from \( u \) to \( v \)), and since \( \epsilon = \frac{\epsilon}{q} \geq \frac{\epsilon}{4q^2} r \).

For the induction step, take edge \( e \in B_i \) of length \( r \) and consider its distance from a disjoint edge \( e' \in B_i \). 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 \( p(e') \). The length of \( p(e) \) is \( r' \), and by construction \( r \leq r' \leq 4r \).

If \( p(e), p(e') \) are disjoint, then by assumption their distance is at least \( \frac{\epsilon}{4q^{2j}} r' \). Since the children of all \( p(e), p(e') \) use new coordinates, the distance from all children of \( p(e) \) to all children of \( p(e') \) is at least the distance from \( p(e) \) to \( p(e') \), that is \( \frac{\epsilon}{4q^{2j}} r' \geq \frac{\epsilon}{4q^{2j}} r \).

Assume then that \( p(e), p(e') \) intersect: Since \( e, e' \in B_i \) are disjoint and \( p(e), p(e') \in B_{i-1} \) are not, at least one of the two parent edges was decomposed into child edges, and by construction the longer edge \( p(e) \) must have been decomposed. So \( r' \geq \frac{\epsilon}{q} \). We consider two cases:

First consider the case where edges \( p(e), p(e') \) intersect in a joint similar to that of edges \( \{a, s\}, \{s, u\} \) or \( \{s, u\}, \{u, t\} \) of the Laakso graph. Then the distance from \( p(e) \) to all disjoint child edges of \( p(e') \) is at
least a quarter of the length of \( p(e') \). Since \( p(e') \) is in the same level as \( p(e) \), its length must be greater than \( \frac{\epsilon r}{4q} \), so a quarter of \( p(e') \)'s length is \( \frac{\epsilon r}{16q} \geq \frac{\epsilon r}{4q^2} \). This also bounds the distance from all children of \( p(e') \) to all disjoint children of \( p(e) \).

Next consider then the case where parent edges \( p(e), p(e') \) meet in a joint such as that formed by edges \( \{s, u\}, \{s, v\} \) of the Laakso graph. In this case the lengths of \( p(e), p(e') \) must be equal. The distance from \( p(e) \) to disjoint child edges of \( p(e') \) is greater than the distance from the (hypothetical) edge \( \{s, t\} \) to all disjoint child edges of \( p(e') \). This in turn is a quarter of the length from \( \{s, t\} \) to \( u; \frac{1}{4} \cdot e|e|' \geq \frac{\epsilon r}{4q^2} r \). This also bounds the distance from all children of \( p(e') \) to all disjoint children of \( p(e) \).

We can then prove the following lemma:

**Lemma 3.** Let \( E \) be a collection of disjoint edges in \( B_i \) with inter-edge distance at most \( 30\epsilon q^4 \). Then \( |E| = O(1) \).

**Proof.** Consider any two edges \( e, e' \in E \), and let \( p(e), p(e') \in B_j \) be their respective ancestors in level \( j = i - 5 \). We will show that \( p(e), p(e') \) must intersect (or be the same edge): Assume that \( p(e) \) is not shorter than \( p(e') \), and let its length be \( r \in (\frac{\epsilon r}{q^2}, q^2] \). By the assertion of the lemma together with \( (11) \), the distance between ancestral edges \( p(e), p(e') \in B_j \) is at most \( 30\epsilon q^4 + 2 \frac{3}{2} \epsilon r = 120\epsilon q^5 e \frac{q^2}{4} + 3\epsilon r < (120\epsilon q^5 + 3\epsilon r) < 3.2\epsilon r \). However, by **Lemma 2**, if \( p(e), p(e') \) are disjoint their distance must be at least \( \frac{\epsilon r}{4q^2} r > 3.7\epsilon r \). So the ancestral edges intersect.

The maximum number of mutually adjacent edges in \( B_j \) is 3, so there are at most 3 ancestral edges in \( B_j \) for all edges in \( E \). Since each edge can beget 6 child edges, \( |E| \leq 3 \cdot 6^5 \).

We can now bound the doubling dimension of \( A_k \): Consider the maximal set \( S \in A_k \) 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 \( |S|^2 \) [26]. Choose \( i \) such that \( \frac{q^i}{4} < \frac{\gamma}{124} < q^i \), and consider graph \( B_i \). Associate each point \( x \in S \) with the closest point \( s(x) \) on its ancestral edge in \( B_i \), and by \( (11) \) the distance of \( x \) from \( s(x) \) is less than \( \frac{3}{2} \epsilon q^i \leq \frac{\gamma}{2} \). Thus the projected points \( \{s(x)\}_{x \in S} \) have inter-point distance in the range \( (\frac{\gamma}{2}, \frac{5\gamma}{2}) \). This distance range implies that any single edge of \( B_i \) can have at most 5 points projected upon it. By **Lemma 3**, the number of edges in \( B_i \) with interpoint distance at most \( \frac{5\gamma}{2} \leq 30\epsilon q^4 \) is bound by a constant, so \( |S| = O(1) \).

## 5 The \( \ell_2 \) case admits dimension reduction

In the previous sections we demonstrated that our construction for \( \ell_p \) rules out dimension reduction for \( p > 2 \). Here we show that our proof fails for \( p = 2 \), and that the \( \ell_2 \) graph admits dimension reduction with low distortion. For simplicity, we will present a constant factor distortion embedding into the plane. Similar techniques can be used to embed into a fixed dimensional \( \ell_2 \) with arbitrarily low distortion.

Our embedding \( h \) into the plane is the natural one, a simple drawing of the graph in two dimensions, which is done as follows: in the construction of **Section 2**, we replace the basis vector \( e_i \) with a unit vector orthogonal to the base edge \( \{a, b\} \) (chosen arbitrarily from the two orthogonal unit vectors in the plane). Since \( \ell_2 \) is invariant under rotations, the embedding \( h \) is isometric on individual instances of the Laakso graph: there is no distortion on any pair of points in the set consisting of any edge in \( B_i \) along with its six child edges in \( B_{i+1} \) (recall the definition of \( B_i \) from **Section 4**). In **Lemma 5**, we prove that \( h \) has constant distortion.

Before attempting the proof, we introduce another embedding from every point in \( B_j \) to a point on its ancestor edge in some previous level, as follows. For all graphs \( B_j, j = 0, \ldots, k \), recursively define
the mapping \( g_i^{(j)} : B_j \to B_i \) for \( i \leq j \) as follows. For any point \( x \) on an edge of \( B_j \) let \( g_j^{(j)}(x) = x \), and let \( g_{j-1}^{(j)}(x) \) be the closest point to \( x \) on its parent edge in \( B_{j-1} \). Finally, when \( i \leq j - 2 \), let \( g_i^{(j)}(x) = g_i^{(i+1)}(g_{i+1}^{(j)}(x)) \). We omit the superscript when \( j \) is clear from the context. Intuitively, \( g_i(x) \) is obtained by following a sequence of projections into parent edges. Observe that \( g_i(x) \) indeed lies on an edge of \( B_i \).

We have the following lemma:

**Lemma 4.** Given a point \( x \in B_j \), let \( g_i(x) \) be found on an edge \( e = \{a, b\} \in B_i \) of length \( r \). Then

1. If \( i < j \), \( \|x - g_{i+1}(x)\|_2 < 0.35er \).
2. If \( i < j \), \( \|h(x) - h(g_{i+1}(x))\|_2 < 0.35er \).
3. \( \|x - g_i(x)\|_2 \leq 3.4 \epsilon \cdot \min\{\|a - g_i(x)\|_2, \|b - g_i(x)\|_2\} \).
4. \( \|h(x) - h(g_i(x))\|_2 \leq 3.4 \epsilon \cdot \min\{\|a - g_i(x)\|_2, \|b - g_i(x)\|_2\} \).

**Proof.** For the proof of the first item, recall that \( q = ((1/4^2 + \epsilon^2)^{1/2}. By a calculation as in (11) and noting that any child edge of \( \{a, b\} \) in level \( i + 1 \) has length at most \( qr \), we obtain that

\[
\|x - g_{i+1}(x)\|_2 < 1.35\epsilon qr < 0.35er,
\]

where we have assumed that \( \epsilon \) (and therefore \( q \)) is sufficiently small. The proof of the second item is similar, using also the fact that \( h \) is isometric on the Laakso graph.

The proof of the third item proceeds by induction. The base case is when \( i = j \), and since \( g_j(x) = x \) the statement is trivial. Suppose then that \( i < j \). We consider two cases: (a) If \( g_{i+1}(x) \) lies on one of the child edges \( \{s, u\}, \{s, v\}, \{u, t\}, \{v, t\} \in B_{i+1} \) of \( \{a, b\} \in B_i \); Then \( g_i(x) = g_i^{(i+1)}(g_{i+1}^{(j)}(x)) \) is in the middle half of parent edge \( \{a, b\} \in B_i \), and so \( \frac{1}{2} \leq \min\{\|a - g_i(x)\|_2, \|b - g_i(x)\|_2\} \leq \frac{1}{2} \). Assume without loss of generality that the minimum is attained for \( a \); then if \( \|a - g_i(x)\|_2 = (1 + \alpha) \cdot \frac{1}{2} \) for some \( \alpha \in [0, 1] \) we have that

\[
\|g_{i+1}(x) - g_i(x)\|_2 = \epsilon \alpha \cdot \frac{1}{2} \leq 2 \epsilon \cdot \frac{1}{4} \cdot r = 2 \epsilon \cdot \min\{\|a - g_i(x)\|_2, \|b - g_i(x)\|_2\}.
\]

Now recalling the first item, we obtain that

\[
\|x - g_{i+1}(x)\|_2 < 0.35er \leq 1.4 \epsilon \min\{\|a - g_i(x)\|_2, \|b - g_i(x)\|_2\}.
\]

It follows that

\[
\|x - g_i(x)\|_2 \leq \|x - g_{i+1}(x)\|_2 + \|g_{i+1}(x) - g_i(x)\|_2 \leq 3.4 \epsilon \min\{\|a - g_i(x)\|_2, \|b - g_i(x)\|_2\}.
\]

(b) If \( g_{i+1}(x) \) lies on one of the child edges \( \{a, s\}, \{t, b\} \in B_{i+1} \) of \( \{a, b\} \), then \( g_i(x) = g_{i+1}(x) \), and thus by the inductive assumption the assertion holds for the child edge, and therefore on \( \{a, b\} \) as well.

The proof of the fourth item is similar to that of the third, using the fact that \( h \) is isometric on the Laakso graph.

**Lemma 5.** The embedding \( h \) has distortion \( O(1) \).
Proof. Consider two points \( x, y \in B_k \). Let integer \( i \) be the largest value for which \( g_i(x), g_i(y) \) map to the same edge \( \{a, b\} \) of length \( r \). If \( i \geq k - 1 \), then the isometry of \( h \) on individual instances of the Laakso graph implies that \( h \) does not distort the distance between \( x \) and \( y \). Assume then that \( i < k - 1 \), and consider respective edges \( e, e' \in B_{i+1} \) onto which \( x, y \) are mapped by \( g_{i+1} \). Suppose that \( e \) and \( e' \) do not intersect: The closest pairs of non-intersecting edges are \( \{s, u\}, \{v, t\} \) (and \( \{s, v\}, \{u, t\} \)). Let \( m \) be the midpoint of edge \( \{a, b\} \); the distance from \( m \) to \( \{s, u\} \) is the altitude to the hypotenuse of right triangle \( \Delta smu \), equal to

\[
\frac{\|s - m\|_2 \cdot \|u - m\|_2}{\|s - u\|_2} = \frac{er/4}{\sqrt{1/4}^2 + \epsilon^2} = \frac{er}{\sqrt{1 + (4\epsilon)^2}},
\]

and the distance from \( \{s, u\} \) to \( \{v, t\} \) (and \( \{s, v\} \) to \( \{u, t\} \)) is twice this. So when \( \epsilon \) is small enough,

\[
\|g_{i+1}(x) - g_{i+1}(y)\|_2 \geq \frac{2er}{\sqrt{1 + (4\epsilon)^2}} > \frac{3\epsilon r}{2}. \tag{13}
\]

Invoking the triangle inequality and the first assertion of Lemma 4, we have that

\[
\|x - y\|_2 \geq \|g_{i+1}(x) - g_{i+1}(y)\|_2 - \|x - g_{i+1}(x)\|_2 - \|y - g_{i+1}(y)\|_2 \geq (1 - \frac{2 \cdot 0.35}{3/2})\|g_{i+1}(x) - g_{i+1}(y)\|_2 = \left(1 - \frac{7}{15}\right)\|g_{i+1}(x) - g_{i+1}(y)\|_2, \tag{14}
\]

and similarly

\[
\|x - y\|_2 \leq \|g_{i+1}(x) - g_{i+1}(y)\|_2 + \|x - g_{i+1}(x)\|_2 + \|y - g_{i+1}(y)\|_2 \leq (1 + \frac{2 \cdot 0.35}{3/2})\|g_{i+1}(x) - g_{i+1}(y)\|_2 = \left(1 + \frac{7}{15}\right)\|g_{i+1}(x) - g_{i+1}(y)\|_2, \tag{15}
\]

Using the second assertion of Lemma 4 along with the triangle inequality and the isometry of \( h \) on the Laakso graph, it follows that

\[
\|h(x) - h(y)\|_2 \geq \|h(g_{i+1}(x)) - h(g_{i+1}(y))\|_2 - \|h(g_{i+1}(x)) - h(x)\|_2 - \|h(g_{i+1}(y)) - h(y)\|_2 \geq \left(1 - \frac{7}{15}\right)\|g_{i+1}(x) - g_{i+1}(y)\|_2 \geq \frac{1 - 7/15}{1 + 7/15}\|x - y\|_2, \tag{15}
\]

and also

\[
\|h(x) - h(y)\|_2 \leq \|h(g_{i+1}(x)) - h(g_{i+1}(y))\|_2 + \|h(g_{i+1}(x)) - h(x)\|_2 + \|h(g_{i+1}(y)) - h(y)\|_2 \leq \left(1 + \frac{7}{15}\right)\|g_{i+1}(x) - g_{i+1}(y)\|_2 \leq \frac{1 + 7/15}{1 - 7/15}\|x - y\|_2, \tag{14}
\]

\[\square\]
which together imply constant distortion.

Suppose then that \( e, e' \) intersect; the hardest case will be when \( e, e' \) are the edges \( \{s, u\}, \{s, v\} \) of the Laakso instance. The case that they are the edges \( \{u, t\}, \{v, t\} \) is symmetric, and we leave the remaining easier cases to the reader. Note that for any point pair \( w \) and \( \epsilon \), assertion of Lemma 4, and applying (16) with \( \epsilon \), the edge connecting these points intersects the (hypothetical) segment \( \{s, t\} \). The distance from \( w \) (respectively \( w' \)) to \( \{s, t\} \) is exactly

\[
\|w - w'\|_2 \geq \frac{\epsilon}{q} (\|w - s\|_2 + \|w' - s\|_2),
\]

where equality holds whenever \( w, w' \) are equidistant from \( s \). Invoking the triangle inequality, the third assertion of Lemma 4, and applying (16) with \( w = g_{i+1}(x) \) and \( w' = g_{i+1}(y) \), we have that

\[
\|x - y\|_2 \geq \|g_{i+1}(x) - g_{i+1}(y)\|_2 - \|x - g_{i+1}(x)\|_2 - \|y - g_{i+1}(y)\|_2 \\
> \|g_{i+1}(x) - g_{i+1}(y)\|_2 - 3.4\epsilon (\|s - g_{i+1}(x)\|_2 + \|s - g_{i+1}(y)\|_2) \\
\geq (1 - 3.4q)\|g_{i+1}(x) - g_{i+1}(y)\|_2,
\]

and similarly

\[
\|x - y\|_2 \leq \|g_{i+1}(x) - g_{i+1}(y)\|_2 + \|x - g_{i+1}(x)\|_2 + \|y - g_{i+1}(y)\|_2 \\
< \|g_{i+1}(x) - g_{i+1}(y)\|_2 + 3.4\epsilon (\|s - g_{i+1}(x)\|_2 + \|s - g_{i+1}(y)\|_2) \\
\leq (1 + 3.4q)\|g_{i+1}(x) - g_{i+1}(y)\|_2.
\]

By a similar calculation, this time with the fourth assertion of Lemma 4, it follows that

\[
\|h(x) - h(y)\|_2 \\
\geq \|h(g_{i+1}(x)) - h(g_{i+1}(y))\|_2 - \|h(x) - h(g_{i+1}(x))\|_2 - \|h(y) - h(g_{i+1}(y))\|_2 \\
= \|g_{i+1}(x) - g_{i+1}(y)\|_2 - \|h(x) - h(g_{i+1}(x))\|_2 - \|h(y) - h(g_{i+1}(y))\|_2 \\
\geq \|g_{i+1}(x) - g_{i+1}(y)\|_2 - 3.4\epsilon (\|s - g_{i+1}(x)\|_2 + \|s - g_{i+1}(y)\|_2) \\
\geq (1 - 3.4q)\|g_{i+1}(x) - g_{i+1}(y)\|_2 \\
\geq \frac{1 - 3.4q}{1 + 3.4q} \|x - y\|_2,
\]

and similarly

\[
\|h(x) - h(y)\|_2 \\
\leq \|h(g_{i+1}(x)) - h(g_{i+1}(y))\|_2 + \|h(x) - h(g_{i+1}(x))\|_2 + \|h(y) - h(g_{i+1}(y))\|_2 \\
= \|g_{i+1}(x) - g_{i+1}(y)\|_2 + \|h(x) - h(g_{i+1}(x))\|_2 + \|h(y) - h(g_{i+1}(y))\|_2 \\
\leq \|g_{i+1}(x) - g_{i+1}(y)\|_2 + 3.4\epsilon (\|s - g_{i+1}(x)\|_2 + \|s - g_{i+1}(y)\|_2) \\
\leq (1 + 3.4q)\|g_{i+1}(x) - g_{i+1}(y)\|_2 \\
\leq \frac{1 + 3.4q}{1 - 3.4q} \|x - y\|_2,
\]

which together imply constant distortion (for sufficiently small \( \epsilon \)).
References


[15] Nader H. Bshouty, Yi Li, and Philip M. Long. Using the doubling dimension to analyze the general-

Symposium on Foundations of Computer Science, FOCS ’08*, pages 761–770, Washington, DC, USA,


[19] Moses Charikar and Amit Sahai. Dimension reduction in the $\ell_1$ norm. In *43rd Symposium on
Foundations of Computer Science (FOCS 2002), 16-19 November 2002, Vancouver, BC, Canada, Pro-


[22] Sorelle A. Friedler and David M. Mount. Approximation algorithm for the kinetic robust k-center


In *COLT*, 2010.

SIAM, 2011.


[27] Lee-Ad Gottlieb and Liam Roditty. Improved algorithms for fully dynamic geometric spanners and
gеometric routing. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algo-
Mathematics.

[28] 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 Com-

[29] Anupam Gupta, Ilan Newman, Yuri Rabinovich, and Alistair Sinclair. Cuts, trees and $l_1$-embeddings


