Two-Sided Kirszbraun Theorem

Arturs Backurs
Toyota Technological Institute at Chicago (TTIC), IL, USA

Sepideh Mahabadi
Toyota Technological Institute at Chicago (TTIC), IL, USA

Konstantin Makarychev
Northwestern University, Evanston, IL, USA

Yury Makarychev
Toyota Technological Institute at Chicago (TTIC), IL, USA

Abstract

In this paper, we prove a two-sided variant of the Kirszbraun theorem. Consider an arbitrary subset X of Euclidean space and its superset Y. Let f be a 1-Lipschitz map from X to $\mathbb{R}^m$. The Kirszbraun theorem states that the map f can be extended to a 1-Lipschitz map $\tilde{f}$ from Y to $\mathbb{R}^m$. While the extension $\tilde{f}$ does not increase distances between points, there is no guarantee that it does not decrease distances significantly. In fact, $\tilde{f}$ may even map distinct points to the same point (that is, it can infinitely decrease some distances). However, we prove that there exists a $(1+\varepsilon)$-Lipschitz outer extension $\tilde{f}: Y \to \mathbb{R}^{m'}$ that does not decrease distances more than "necessary". Namely,

$$\|\tilde{f}(x) - \tilde{f}(y)\| \geq c\sqrt{\varepsilon}\min(\|x-y\|, \inf_{a,b\in X}(\|x-a\| + \|f(a)-f(b)\| + \|b-y\|))$$

for some absolutely constant $c > 0$. This bound is asymptotically optimal, since no L-Lipschitz extension g can have $\|g(x)-g(y)\| > L\min(\|x-y\|, \inf_{a,b\in X}(\|x-a\| + \|f(a)-f(b)\| + \|b-y\|))$ even for a single pair of points x and y.

In some applications, one is interested in the distances $\|\tilde{f}(x) - \tilde{f}(y)\|$ between images of points $x, y \in Y$ rather than in the map $\tilde{f}$ itself. The standard Kirszbraun theorem does not provide any method of computing these distances without computing the entire map $\tilde{f}$ first. In contrast, our theorem provides a simple approximate formula for distances $\|\tilde{f}(x) - \tilde{f}(y)\|$.

2012 ACM Subject Classification Theory of computation → Computational geometry; Mathematics of computing

Keywords and phrases Kirszbraun theorem, Lipschitz map, Outer-extension, Two-sided extension

Digital Object Identifier 10.4230/LIPIcs.SoCG.2021.13

Funding *Arturs Backurs*: Supported in part by NSF award CCF-2006806.
Konstantin Makarychev: Supported in part by NSF awards CCF-1955351 and HDR TRIPODS CCF-1934931.
Yury Makarychev: Supported in part by NSF awards CCF-1718820, CCF-1955173, and HDR TRIPODS CCF-1934843.

1 Introduction

In this paper, we prove a two-sided variant of the Kirszbraun theorem. The Kirszbraun theorem [9] is widely used in high-dimensional geometry and analysis, and has recently found applications in theoretical computer science and machine learning [1, 2, 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 16]. Recall that a function f from a subset S of Euclidean space[1] ℓ_2^n to Euclidean space ℓ_2^m is L-Lipschitz if it increases distances between points by at most a factor of L:

[1] We denote d-dimensional Euclidean space by ℓ_2^d; i.e., ℓ_2^d is $\mathbb{R}^d$ with the standard Euclidean distance.

37th International Symposium on Computational Geometry (SoCG 2021).
Editors: Kevin Buchin and Éric Colin de Verdière; Article No. 13; pp. 13:1–13:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

$\|f(x) - f(y)\| \leq L\|x - y\|$ for all $x, y \in X$. The Lipschitz constant $\|f\|_{Lip}$ of a function f is the minimum number L such that f is L-Lipschitz. Now, consider a subset S of Euclidean space ℓ_2^n, its superset $T \subset \ell_2^n$, and an L-Lipschitz map $f : S \to \mathbb{R}^m$. The Kirszbraun theorem states that f can be extended to a map $\tilde{f} : T \to \mathbb{R}^m$ without increasing its Lipschitz constant. That is, the theorem guarantees that $\tilde{f}$ does not increase distances by more than a factor of L. Can we guarantee that distances $\|\tilde{f}(x) - \tilde{f}(y)\|$ do not *decrease* significantly? The theorem does not provide us with any such guarantee; in fact, $\tilde{f}$ may even map distinct points to the same point and thus contract some distances infinitely.

In this paper, we prove that there exists an extension $\tilde{f}$ that does not decrease distances more than "necessary". What kind of distance contraction is necessary? Consider sets $S \subset T \subset \ell_2^n$ and an L-Lipschitz map $f : S \to \ell_2^m$. Note if f significantly contracts the distance between $a, b \in S$, then so must $\tilde{f}$, since f and $\tilde{f}$ coincide on a and b. Moreover, every L-Lipschitz extension $\tilde{f}$ must map points that are close to a and b to points close to $f(a)$ and $f(b)$, which we assumed are close to each other. More generally, consider arbitrary points $x, y \in T$, $a, b \in S$, and a cL-Lipschitz extension $\tilde{f}$ of f (where $c \geq 1$). Then, we have

$$\begin{aligned}\|\tilde{f}(x) - \tilde{f}(y)\| &\leq \|\tilde{f}(x) - \tilde{f}(a)\| + \|\tilde{f}(a) - \tilde{f}(b)\| + \|\tilde{f}(b) - \tilde{f}(y)\| \\ &\leq cL\|x - a\| + \|f(a) - f(b)\| + cL\|b - y\|. \qquad (1)\end{aligned}$$

Since $\tilde{f}$ is cL-Lipschitz, we also have

$$\|\tilde{f}(x) - \tilde{f}(y)\| \leq cL\|x - y\|. \qquad (2)$$

We see that $\tilde{f}$ must satisfy (1) for all $a, b \in S$ and (2). We restate this condition in the following claim.

▷ Claim 1. Let $S \subset T \subset \ell_2^n$ and f be an L-Lipschitz map from S to ℓ_2^m. Define metric $d_{\mathrm{ub}}(\cdot, \cdot)$ on T as follows.

$$d_{\mathrm{ub}}(x, y) = \min\bigl(L\|x - y\|, \inf_{a,b \in S} (L\|x - a\| + \|f(a) - f(b)\| + L\|b - y\|)\bigr). \qquad (3)$$

Then for every $c \geq 1$ and a cL-Lipschitz extension $\tilde{f} : T \to \ell_2^m$ of f, we have

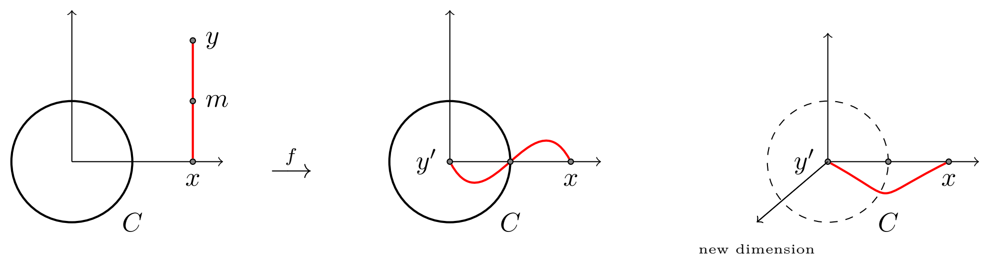
$$\|\tilde{f}(x) - \tilde{f}(y)\| \leq c d_{\mathrm{ub}}(x, y) \quad \text{for all } x, y \in T.$$

Note that $d_{\mathrm{ub}}(\cdot, \cdot)$ satisfies all the axioms of a metric except that $d_{\mathrm{ub}}(x, y)$ may be equal to 0 for $x \neq y$.

Our goal now is to prove a "tight" variant of the Kirszbraun theorem: there exists a Lipschitz extension $\tilde{f}$ of f such that $\|\tilde{f}(x) - \tilde{f}(y)\| \geq \Omega(d_{\mathrm{ub}}(x, y))$ for every x and y.

However, in order to obtain such a result or, for that matter, any non-trivial result, we need to relax two conditions in the Kirszbraun theorem. We use the following example to explain what these conditions are and why we need to relax them.

► **Example 2.** Consider a set S that consists of a circle C around point $(0, 0) \in \ell_2^2$ and two points $x = (2, 0)$ and $y = (2, 2)$.



Define f as follows: f maps each point of C to itself, x to x, and y to $y' = (0,0)$. Let $T = S \cup [x,y]$. It is immediate that f is 1-Lipschitz and $d_{\text{ub}}(u,v) > 0$ for every pair of points u and v. However, observe that the images of C and $[a,b]$ under every Lipschitz extension $\tilde{f} : T \to \ell_2^2$ necessarily intersect. Thus, every Lipschitz extension $\tilde{f}$ infinitely decreases the distance between some pair of points.

To overcome this obstacle, we consider *outer* extensions $\tilde{f}$ of f [14], which are allowed to use additional dimensions/coordinates. As standard, we denote the direct sum of Euclidean or Hilbert spaces ℓ_2^a and ℓ_2^b by $\ell_2^a \oplus \ell_2^b$, which is isometrically isomorphic to ℓ_2^{a+b}.

▶ **Definition 3.** *Let S be a non-empty subset of Euclidean space ℓ_2^n or separable Hilbert space ℓ_2^∞ and T be its superset. Consider a map $f : S \to \ell_2^m$, where m can be finite or infinite. Map $\tilde{f}$ from T to $\ell_2^m \oplus \ell_2^\Delta \simeq \ell^{m'}$ is an outer extension of $f : S \to \ell_2^m$ to $T \supset S$ (where Δ may be finite or infinite, and $m' = m + \Delta$) if it maps every point $x \in S$ to $f(x) \oplus \bar{0}$.*[2] *We will call extensions that do not use extra coordinates proper extensions.*

In Example 2, a Lipschitz *outer* extension $\tilde{f}$ can map segment $[x,y]$ to a curve that starts at x, goes above the plane ℓ_2^2, and then enters y. Map $\tilde{f}$ no longer maps distinct points to the same point. However, the Lipschitz constant of $\tilde{f}$ must be greater than 1. Indeed, let m be the midpoint between x and y. Point m is at distance 1 from each of the points x and y. If $\tilde{f}$ were 1-Lipschitz, then it would map m to a point whose distances to $x = f(x)$ and $y' = f(y)$ would not exceed 1. However, the only such point is the midpoint between x and y', which lies on C. Thus, if $\tilde{f}$ were 1-Lipschitz, it would map distinct points to the same point and thus would contract some distances infinitely. Therefore, given an L-Lipschitz map f, we will construct a $(1+\varepsilon)L$-Lipschitz outer extension $\tilde{f}$ rather than an L-Lipschitz extension. We are ready to state our main result.

▶ **Theorem 4** (Two-sided Kirszbraun Theorem)**.** *Consider $m, n \in \mathbb{Z}_{\geq 1} \cup \{\infty\}$ and $\varepsilon \in [0,1]$. Let $S \subset T$ be non-empty subsets of ℓ_2^n and f be an L-Lipschitz map from S to ℓ_2^m. Let $d_{\text{ub}}(\cdot,\cdot)$ be as defined by formula (3). There exists a $(1+\varepsilon)L$-Lipschitz outer extension $\tilde{f}$ from T to $\ell_2^m \oplus \ell_2^\Delta \simeq \ell_2^{m'}$ such that*

$$\|\tilde{f}(x) - \tilde{f}(y)\| \geq c\sqrt{\varepsilon}\, d_{\text{ub}}(x,y) \quad \textit{for all } x, y \in T.$$

Here, c is an absolute constant. If $|T \setminus S|$ is finite, then $\Delta = O(\log |T \setminus S|)$ and $m' = m + \Delta$; if $|T \setminus S|$ is infinite, then $\Delta = m' = \infty$.

[2] That is, if m is finite, we identify ℓ_2^m with the subspace of $\ell_2^{m'}$ spanned by the first m basis vectors and allow $\tilde{f}$ to take values in $\ell_2^{m'}$.

We discuss some properties of $\tilde{f}$. It is easy to see that Claim 1 applies to both proper and outer extensions. Therefore, $\|\tilde{f}(x) - \tilde{f}(y)\| \le (1+\varepsilon)d_{\mathrm{ub}}(x,y)$. In particular, when $\varepsilon \in (0,1]$ is fixed, distances $\|\tilde{f}(x) - \tilde{f}(y)\|$ and $\Theta(d_{\mathrm{ub}}(x,y))$ are within a constant factor of each other.

$$\frac{\|\tilde{f}(x) - \tilde{f}(y)\|}{d_{\mathrm{ub}}(x,y)} \in [c\sqrt{\varepsilon}, 1+\varepsilon]. \tag{4}$$

Least possible contraction for each pair (x, y). Another property of the map $\tilde{f}$ is that it asymptotically contracts distances less than any other $(c'L)$-Lipschitz proper or outer extension g of f:

$$\|\tilde{f}(x) - \tilde{f}(y)\| \overset{\text{by Theorem 4}}{\ge} c\sqrt{\varepsilon} d_{\mathrm{ub}}(x,y) \overset{\text{by Claim 1}}{\ge} \frac{c\sqrt{\varepsilon}}{c'}\|g(x) - g(y)\| \quad \text{for all } x, y \in T.$$

Easy to compute distances. For finite subsets S and T, we can efficiently compute extensions – whose existence is guaranteed by the standard Kirszbraun theorem and Theorem 4 – using semidefinite programming (SDP).[3] However, in many proofs and applications we need to know only distances $\|\tilde{f}(x) - \tilde{f}(y)\|$ and not the map $\tilde{f}$ itself. The Kirszbraun theorem does not provide any method for obtaining the distances other than computing the entire map $\tilde{f}$, using SDP or MWU, and then directly computing $\|\tilde{f}(x) - \tilde{f}(y)\|$. In contrast, Theorem 4 provides a simple formula (4) for approximately computing all pairwise distances.

Optimal parameters

The following theorem shows that the parameters in Theorem 4 cannot be significantly improved.

▶ **Theorem 5.** *The following items hold.*

- *There exist finite sets $S \subset \ell_2^2$ and $T = S \cup \{z_1, z_2\} \subset \ell_2^2$ and a 1-Lipschitz function $f : S \to \ell_2^2$ such that the following is true. For every $\varepsilon \in (0,1]$ and a $(1+\varepsilon)$-Lipschitz extension $\tilde{f}$,*

$$\|\tilde{f}(z_1) - \tilde{f}(z_2)\| \le O(\sqrt{\varepsilon}\, d_{\mathrm{ub}}(z_1, z_2)).$$

- *For every $m, n, N \ge 1$, there exist finite sets $S \subset T \subset \ell_2^n$ with $|T \backslash S| = N$ and a 1-Lipschitz map $f : S \to \ell_2^m$ such that the following is true. For every $c > 0$, if $\tilde{f} : \ell_2^n \to \ell_2^{m'}$ is an L-outer extension and $\|\tilde{f}(x) - \tilde{f}(y)\| \ge c\, d_{\mathrm{ub}}(x,y)$ for every $x, y \in T$, then $m' \ge c' \log_e N$ where $c' = 1/(\log_e(L/c+1))$.*
- *For every $m, n \ge 1$, there exist infinite sets $S \subset T \subset \ell_2^n$ and a 1-Lipschitz map $f : S \to \ell_2^m$ such that for every Lipschitz extension $\tilde{f} : T \to \ell_2^{m'}$ satisfying $\|\tilde{f}(x) - \tilde{f}(y)\| \ge c\, d_{\mathrm{ub}}(x,y)$ (for every $x, y \in T$ and some $c > 0$), we have $m' = \infty$.*

[3] Standard Kirszbraun extensions can be also computed using quadratically constrained quadratic programming (QCQP) or the multiplicative weight update method (MWU) [3].

1.1 Applications

Next we discuss two applications of our result.

Updating Euclidean metric

Similarity information in data sets is often encoded with Euclidean distance: two data points are similar if and only if they are close to each other (one can store either the distance itself or an embedding of points in Euclidean space). Consider the scenario where we initially have a data set X and some information about objects in X. Based on this information, we compute a Euclidean distance d_X on X. Then, we get an updated information about some subset of points Y. Using this information, we compute a new Euclidean distance d_Y on Y. Now we want to update d_X with new distances from d_Y. A natural way to do this is to first "combine" the metrics into a distance function $d_c(\cdot,\cdot)$

$$d_c = \begin{cases} d_Y(x,y), & \text{if } x,y \in Y \\ d_X(x,y), & \text{otherwise} \end{cases}. \tag{5}$$

Since d_c is not necessarily a metric, we consider the metric closure d_u of d_c (recall that by definition $d_u(x,y)$ is the length of the shortest path between x and y in the complete graph on X with edge lengths $d_c(x,y)$; see Definition 10). Metric d_u is our updated metric. We want d_u to be Euclidean (as d_X and d_Y are), but, in general, it does not have to be Euclidean.

▶ **Definition 6.** *A map f from a metric space (U, d_U) to a metric space (V, d_V) has distortion at most $D \geq 1$ if for some $c > 0$ and every $x, y \in U$,*

$$c d_U(x,y) \leq d_V(f(x), f(y)) \leq cDd_U(x,y).$$

We say that a metric space (U, d_U) is D-Euclidean or that it embeds into Euclidean space with distortion at most D, if there is an embedding $f : U \to \ell_2^m$ (for some $m \in \mathbb{N} \cup \infty$) with distortion at most D.

We provide a sufficient condition when d_u is Euclidean.

▶ **Theorem 7.** *I. Consider a finite D_X-Euclidean metric space (X, d_X). Let Y be a subset of X and d_Y be a D_Y-Euclidean metric on Y. Assume that $d_Y(x,y) \leq Cd_X(x,y)$ for all $x, y \in Y$ (where $C \geq 1$). Then the updated metric d_u (defined as the closure of d_c; see (5)) is $O(CD_XD_Y)$-Euclidean.*

II. The requirement that $d_Y(x,y) \leq Cd_X(x,y)$ in item I is necessary. For every N, there exist a 1-Euclidean metric space (X, d_X) on at most N points, $Y \subset X$, and 1-Euclidean metric d_Y on Y such that every embedding of the updated metric d_u into ℓ_2 requires distortion at least $\Omega(\log N)$.

Bi-Lipschitz extension

The bi-Lipschitz constant, i.e., distortion, of a map $f\colon X \to Y$ is the minimum D such that for some $\lambda > 0$ and every $x, y \in X$, $\lambda \cdot d_X(x,y) \leq d_Y(f(x), f(y)) \leq \lambda \cdot D \cdot d_X(x,y)$. Recall that the Bi-Lipschitz map, is a map with a bounded distortion. Recently, Mahabadi, Makarychev, Makarychev, and Razenshteyn [14] proved a bi-Lipschitz variant of the Kirszbraun theorem and showed its applications to prioritized dimension reductions.

▶ **Theorem 8** (Kirszbraun theorem for bi-Lipschitz maps [14]). *Consider $S \subset T \subset \ell_2^n$ and a map $f : S \to \ell_2^m$ with distortion at most D. There exists an outer extension $\tilde{f} : T \to \ell_2^{m'}$ with distortion at most $O(D)$, where $m' = m + n$.*

This theorem follows immediately from Theorem 4 with the caveat that we do not get any bound on m' in terms of m and n (in particular, m' might be significantly greater than $m+n$). Indeed, we can assume without loss of generality that $\|f\|_{Lip} = 1$ and $\|f(x) - f(y)\| \geq \|x - y\|/D$ for all $x, y \in S$. Then $d_{\mathrm{ub}}(x, y) \geq \|x - y\|/D$. Thus for the outer extension $\tilde{f}$ from Theorem 4 (say with $\varepsilon = 1/2$), we have $\Omega(\|x - y\|/D) \leq \|\tilde{f}(x) - \tilde{f}(y)\| \leq {}^3\!/_2\|x - y\|$; that is, $\tilde{f}$ has distortion at most $O(D)$.

Summary

In this paper, we prove a two-sided variant of the Kirszbraun theorem. The theorem guarantees that there is a Lipschitz outer extension $\tilde{f}$ that contracts the distance of every pair of points less than any other extension (up to a constant factor) and that has asymptotically optimal parameters. Unlike the standard Kirszbraun theorem, our theorem provides a simple approximate formula for distances $\|\tilde{f}(x) - \tilde{f}(y)\|$. Additionally, we show an application of our theorem to the Euclidean metric update problem.

Organization. In Section 2, we introduce some notation and relevant results as well as formally state the Kirszbraun theorem. In Section 3, we prove Theorem 4. In Section 4, we prove Theorem 5. Finally, in Section 5, we prove Theorem 7.

2 Preliminaries

In this paper, ℓ_2^n denotes the n-dimensional Euclidean space equipped with the standard Euclidean norm $\|\cdot\|$, when $n < \infty$; $\ell_2^\infty = \ell_2$ denotes the infinite dimensional separable Hilbert space. For $m < m'$, we identify ℓ_2^m with the m-dimensional subspace of $\ell_2^{m'}$ spanned by the first m standard basis vectors (in other words, we identify vectors $(x_1, \dots, x_m) \in \ell_2^m$ and $(x_1, \dots, x_m, 0, \dots, 0) \in \ell_2^{m'}$).

We will need the following theorem proved by Mendel and Naor [17].

▶ **Theorem 9** (Lemma 5.2 in [17], restated). *Consider space ℓ_2^n (where n is finite or infinite). For every $r > 0$, there exists a map $\psi_r : \ell_2^n \to \ell_2$ such that*[4]

$$\sqrt{\frac{e-1}{e}}\min(\|x - y\|, \sqrt{2}\,r) \leq \|\psi_r(x) - \psi_r(y)\| \leq \min(\|x - y\|, \sqrt{2}\,r)$$
$$\|\psi_r(x)\| = r$$

for all $x, y \in \ell_2^n$.

We will consider maps h from $\mathbb{R}$ to Hilbert space ℓ_2 equipped with the L_2 norm:

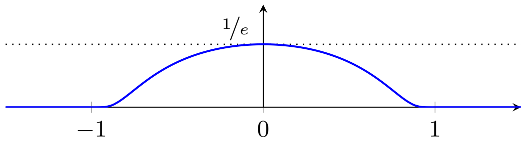
$$\|h\|_{L_2}^2 = \int_{-\infty}^{+\infty} \|h(t)\|^2 dt.$$

[4] The specific constants $\frac{e-1}{e}$ and $\sqrt{2}$ do not appear in the statement of Lemma 5.2 in [17], but can be easily deduced from the proof. Note that there is a typo on the last line of the proof of Lemma 5.2 in [17]: D should be replaced with $\sqrt{2}D$ both in the lower and upper bounds for $\|F(x) - F(y)\|_2$.

As standard, we denote the set of all such functions whose L_2-norm is finite by $L_2(\mathbb{R}, \ell_2)$. We note that Hilbert space $L_2(\mathbb{R}, \ell_2)$ is isometrically isomorphic to ℓ_2.

In the proof, we will use a "bump function" λ:

$$\lambda(t) = \begin{cases} e^{-\frac{1}{1-t^2}}, & \text{if } t \in (-1,1) \\ 0, & \text{otherwise} \end{cases}$$



We will need the following easily verifiable properties of $\lambda(t)$. Function λ is zero outside of $(-1,1)$. It is non-negative and upper bounded by $1/e$. Function λ is everywhere differentiable, and its derivative $\lambda'(t)$ is bounded by 1 in absolute value. Finally, $\lambda(t) > 1/4$ when $t \in (-1/2, 1/2)$ and $\lambda(t) > 1/7$ when $t \in (-2/3, 2/3)$.

▶ **Definition 10** (Metric Closure). *Consider a finite set of points X. Let d be a distance function on X that does not necessarily satisfy the triangle inequality (we do assume that for all $x, y \in X$: (i) $d(x,y) = 0$ if and only if $x = y$ and (ii) $d(x,y) = d(y,x)$). Denote the complete graph on X with edge lengths $d(\cdot,\cdot)$ by $K(X,d)$. The metric closure of d is the shortest path distance in $K(X,d)$.*

▶ **Theorem 11** (Kirszbraun Extension Theorem). *Consider a subset S of Euclidean space ℓ_2^n, its superset $T \subset \ell_2^n$, and an L-Lipschitz map $f : S \to \ell_2^m$. Then there exists an L-Lipschitz extension $\tilde{f} : T \to \ell_2^m$. Dimensions m and n can be finite or infinite.*

3 Two-sided Kirszbraun

In this section, we prove Theorem 4. We start with proving a simple geometric inequality (Lemma 12) and then the main lemma (Lemma 13). Theorem 4 will easily follow from Lemma 13.

Without loss of generality, we assume that set S is closed. If it is not, we let S' be the closure of S and extend f continuously to S'; then we apply the theorem to set S'. If $\varepsilon = 0$, the theorem immediately follows from the standard Kirszbraun theorem, so we assume that $\varepsilon > 0$. Let $R_x = d(x,S) = \min_{y \in S} \|x - y\|$.

▶ **Lemma 12.** *Let $u, v \in \ell_2$ be two vectors of length r and $a \geq b \geq 0$. Then*

$$\max((a-b)r, b\|u-v\|) \leq \|au - bv\| \leq (a-b)r + b\|u-v\|.$$

Proof. First, by the triangle inequality, $\|au - bv\| \leq \|au - bu\| + \|bu - bv\| = (a-b)\|u\| + b\|u-v\| = (a-b)r + b\|u-v\|$. Then $\|au - bv\| \geq \|au\| - \|bv\| = (a-b)r$. Finally, observe that $\langle u - v, u\rangle = r^2 - r^2\cos\alpha \geq 0$, where α is the angle between u and v. Therefore, $\langle bu - bv, au - bu\rangle \geq 0$. We have $\|au - bv\|^2 = \|(bu - bv) + (au - bu)\|^2 = \|bu - bv\|^2 + \|au - bu\|^2 + 2\langle bu - bv, au - bu\rangle \geq \|bu - bv\|^2$. ◀

▶ **Lemma 13.** *There is a map $h : \ell_2^n \to L_2(\mathbb{R}, \ell_2)$ such that*
1. *h maps S to 0,*
2. *$\|h(x) - h(y)\|_{L_2} = \Theta(\min(\|x - y\|, R_x + R_y))$ for all $x, y \in \ell_2^n$.*

Proof. Let ψ be as in Theorem 9 and λ be the bump function defined in Section 2. Define $h : \ell_2^n \to L_2(\mathbb{R}, \ell_2)$ as follows:

$$h(x)(t) = \lambda(\ln R_x - t)\psi_{e^t}(x).$$

Here, we assume that $\ln 0 = -\infty$ and $\lambda(-\infty) = 0$; in other words, $h(x)(t) = 0$ if $R_x = 0$.

Let $I_x = (\ln R_x - 1, \ln R_x + 1)$ for $x \notin S$ and $I_x = \varnothing$ for $x \in S$. We have the following

- $h(x)(t) = 0$ when $t \notin I_x$ (since λ is supported on $(-1, 1)$).
- $\|h(x)(t)\| = \lambda(\ln R_x - t)\|\psi_{e^t}(x)\| = \lambda(\ln R_x - t)e^t < eR_x$ when $t \in I_x$ (since $\|\lambda\|_\infty < 1$ and $\|\psi_{e^t}(x)\| = e^t$).

We verify the first condition: If $x \in S$, then $R_x = 0$ and thus $h(x) = 0$, as required. Now we verify the second condition. Below, we assume without loss of generality that $R_x \geq R_y$.

First, we prove the desired upper bound on $\|h(x) - h(y)\|_{L_2}$ for $x, y \in \ell_2^n$. Note that if $t \notin I_x \cup I_y$, then $\|h(x)(t) - h(y)(t)\| = 0$. Further, the measure of $I_x \cup I_y$ is at most 4. For $t \in I_x \cup I_y$, $\|h(x)(t) - h(y)(t)\| \leq \|h(x)(t)\| + \|h(y)(t)\| \leq e(R_x + R_y)$. Thus,

$$\|h(x) - h(y)\|_{L_2}^2 = \int_{-\infty}^{+\infty} \|h(x)(t) - h(y)(t)\|^2 dt = \int_{I_x \cup I_y} \|h(x)(t) - h(y)(t)\|^2 dt \leq 4e^2(R_x + R_y)^2.$$

We have proved that $\|h(x) - h(y)\|_{L_2} = O(R_x + R_y)$.

Now we show that $\|h(x) - h(y)\|_{L_2} = O(\|x - y\|)$. Note that if $R_x > 2R_y$, then $\|x - y\| \geq R_x - R_y > (R_x + R_y)/3$ and therefore $\|h(x) - h(y)\|_{L_2} \leq O(R_x + R_y) \leq O(\|x - y\|)$, and we are done. So we assume that $R_x \leq 2R_y$. From Lemma 12, we get

$$\begin{aligned}\|h(x)(t) - h(y)(t)\| &= \|\lambda(\ln R_x - t)\psi_{e^t}(x) - \lambda(\ln R_y - t)\psi_{e^t}(y)\| \\ &\leq |\lambda(\ln R_x - t) - \lambda(\ln R_y - t)|e^t + \min(\lambda(\ln R_x - t), \lambda(\ln R_y - t))\|\psi_{e^t}(x) - \psi_{e^t}(y)\|.\end{aligned}$$

We upper bound the first term using the Mean Value Theorem. Using that $\|\lambda'\|_\infty \leq 1$ and $\ln a \leq a - 1$ for every $a \in \mathbb{R}$, we get that for some $\xi \in (\ln R_y - t, \ln R_x - t)$,

$$|\lambda(\ln R_x - t) - \lambda(\ln R_y - t)|e^t = |\lambda'(\xi)| \cdot |(\ln R_x - t) - (\ln R_y - t)|e^t \leq |\ln {}^{R_x}/_{R_y}|e^t \leq \frac{R_x - R_y}{R_y} e^t.$$

Now we upper bound the second term. By Theorem 9,

$$\min(\lambda(\ln R_x - t), \lambda(\ln R_y - t))\|\psi_{e^t}(x) - \psi_{e^t}(y)\| \leq O(\min(e^t, \|x - y\|)) \leq O(\|x - y\|).$$

We get

$$\begin{aligned}\|h(x) - h(y)\|_{L_2}^2 &= \int_{I_x \cup I_y} \|h(x)(t) - h(y)(t)\|^2 dt \leq 4 \left(\max_{t \in I_x \cup I_y} \left(\frac{R_x - R_y}{R_y} e^t \right) + O(\|x - y\|) \right)^2 \\ &\leq O\left(\frac{R_x - R_y}{R_y} \cdot R_x + \|x - y\| \right)^2 \leq O\left(2(R_x - R_y) + \|x - y\|\right)^2.\end{aligned}$$

Since $R_x - R_y \leq \|x - y\|$, we get that $\|h(x) - h(y)\|_{L_2} = O(\|x - y\|)$.

We have proved the desired upper bound on $\|h(x) - h(y)\|_{L_2}$. Now we prove the lower bound. Consider two cases.

Case 1: $R_x \geq eR_y$. In this case, $\ln R_x \geq \ln R_y + 1$. Let $I_x^+ = (\ln R_x, \ln R_x + 1/2)$. Note that I_x^+ does not intersect I_y and thus $h(y)(t) = 0$ for $t \in I_x^+$. Also, since $\lambda(t) \geq 1/4$ on $[-1/2, 1/2]$, we have

$$\|h(y)(t)\|_2 = \lambda(\ln R_x - t)\|\psi_{e^t}(x)\| \geq {}^{e^t}/_4 \geq R_x/4$$

for $t \in I_x^+$. We have,

$$\|h(x) - h(y)\|_{L_2}^2 \geq \int_{I_x^+} \|h(x)(t) - h(y)(t)\|^2 dt = \int_{I_x^+} \|h(x)(t)\|^2 dt \geq \frac{1}{2}\left(\frac{R_x}{4}\right)^2.$$

We conclude that $\|h(x) - h(y)\|_{L_2} = \Omega(R_x + R_y)$.

Case 2: $R_x < eR_y$. In this case, $\ln R_y \le \ln R_x < \ln R_y + 1$. Let $J = (\ln R_y + 1/3, \ln R_y + 2/3)$. Then for $t \in J$, we have $\ln R_y - t \in (-2/3, -1/3)$ and $\ln R_x - t \in (-2/3, 2/3)$. Therefore, $\lambda(\ln R_x - t) \ge 1/7$ and $\lambda(\ln R_y - t) \ge 1/7$. By Lemma 12, we have for $t \in J$,

$$\begin{aligned}\|h(x)(t) - h(y)(t)\| &= \|\lambda(\ln R_x - t)\psi_{e^t}(x) - \lambda(\ln R_y - t)\psi_{e^t}(y)\| \\ &\ge \frac{1}{7}\|\psi_{e^t}(x) - \psi_{e^t}(y)\| \ge \Omega(\min(e^t, \|x-y\|)) \\ &\ge \Omega(\min(R_x + R_y, \|x-y\|))\end{aligned}$$

Using that the length of segment J is $1/3$, we get

$$\|h(x) - h(y)\|_{L_2}^2 \ge \int_J \|h(x)(t) - h(y)(t)\|^2 dt \ge \Omega(\min(R_x + R_y, \|x-y\|))^2.$$

This concludes the proof of the lower bound and the lemma. ◀

Since $L_2(\mathbb{R}, \ell_2)$ is isometrically isomorphic to ℓ_2, we get the following corollary.

▶ **Corollary 14.** *Let $T \subset \ell_2^n$. There is a map $h : T \to \ell_2^\Delta$, where $\Delta = O(\log |T \setminus S|)$ (Δ is infinite if $|T \setminus S|$ is infinite) such that*

1. *h maps S to 0,*
2. *$c \min(\|x-y\|, R_x + R_y) \le \|h(x) - h(y)\| \le \min(\|x-y\|, R_x + R_y)$ for all $x, y \in \ell_2^n$ and some absolute constant $c > 0$.*

Proof. We start with a map $h_0 : \ell_2^n \to L(\mathbb{R}, \ell_2)$ from Lemma 13. Since $L_2(\mathbb{R}, \ell_2)$ is isometrically isomorphic to ℓ_2, there is a map $h_1 : \ell_2^n \to \ell_2$ satisfying the conditions in Lemma 13. Now if $T \setminus S$ is finite, we apply Johnson-Lindenstrauss dimension reduction π_{JL} with distortion at most $3/2$ to $h_1(T)$ [8]. Note that h_1 maps all points in S to 0; hence, $|h_1(T)| \le 1 + |h_1(T \setminus S)| \le 1 + |T \setminus S|$. Therefore, π_{JL} maps $h_1(T \setminus S)$ to ℓ_2^Δ with $\Delta = O(\log |T \setminus S|)$. Without loss of generality, we assume that $\pi_{JL}(0) = 0$. We get map $h_2 = \pi_{JL} \circ h_1$. If $T \setminus S$ is infinite, we simply let $h_2 = h_1$. Finally, we rescale h_2 so that the obtained map h satisfies the desired inequality:

$$c \min(\|x-y\|, R_x + R_y) \le \|h(x) - h(y)\| \le \min(\|x-y\|, R_x + R_y).$$ ◀

We are ready to prove Theorem 4. Let g be the standard Kirszbraun extension of f (Theorem 11) and h be the map provided by Corollary 14. Note that the second condition in Corollary 14 guarantees that h is 1-Lipschitz. Define $\tilde{f}(x) = g(x) \oplus (\sqrt{\varepsilon}L)h(x) \in \ell_2^m \oplus \ell_2^\Delta$.

For $x \in S$, we have $\tilde{f}(x) = f(x) \oplus 0$; thus $\tilde{f}$ is an outer extension of f. Since g is L-Lipschitz and h is 1-Lipschitz,

$$\|\tilde{f}\|_{Lip} \le \sqrt{\|g\|_{Lip}^2 + (\sqrt{\varepsilon}L)^2 \|h\|_{Lip}^2} \le \sqrt{1+\varepsilon}\, L \le (1+\varepsilon)L.$$

Finally, consider arbitrary $x, y \in \ell_2^n$. If n is finite, let x' and y' be points in S closest to x and y, respectively (such points exist, since S is closed). Then $\|x - x'\| = R_x$ and $\|y - y'\| = R_y$. If n is infinite, let x' and y' be points in S such that $\|x - x'\| \le 2R_x$ and $\|y - y'\| \le 2R_y$.

Note that $\|g(x) - g(y)\| \ge \|f(x') - f(y')\| - \|g(x) - f(x')\| - \|g(y) - f(y')\| \ge \|f(x') - f(y')\| - 2L(R_x + R_y)$. We have,

$$\begin{aligned}\|\tilde{f}(x) - \tilde{f}(y)\| &= \sqrt{\|g(x) - g(y)\|^2 + \varepsilon L^2 \|h(x) - h(y)\|^2} \\ &\ge \max(\|g(x) - g(y)\|, \sqrt{\varepsilon}L\|h(x) - h(y)\|) \\ &\ge \max(\|f(x') - f(y')\| - 2L(R_x + R_y), c\sqrt{\varepsilon}L \min(\|x-y\|, R_x + R_y)).\end{aligned}$$

If $\|x-y\| \le R_x + R_y$, then we get that $\|\tilde{f}(x) - \tilde{f}(y)\| \ge c\sqrt{\varepsilon}\, L\|x-y\| \ge c\sqrt{\varepsilon}\, d_{\mathrm{ub}}(x,y)$ and we are done. So we assume now that $\|x-y\| \ge R_x + R_y$. Note that

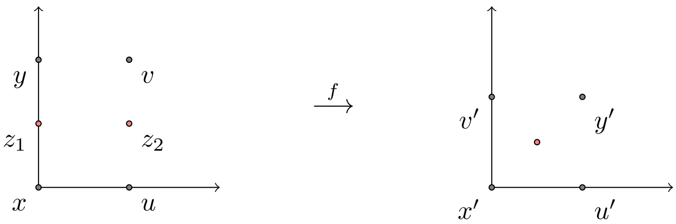
$$\begin{aligned} d_{\mathrm{ub}}(x,y) &\le \inf_{a,b\in S}(L\|x-a\| + \|f(a)-f(b)\| + L\|x-a\|) \\ &\le L\|x-x'\| + \|f(x')-f(y')\| + L\|y'-y\| \\ &\le \|f(x')-f(y')\| + 2L(R_x+R_y). \end{aligned}$$

We have

$$\begin{aligned} \|\tilde{f}(x)-\tilde{f}(y)\| &\ge \max(\|f(x')-f(y')\| - 2L(R_x+R_y), c\sqrt{\varepsilon}L(R_x+R_y)) \\ &\ge \frac{(c\sqrt{\varepsilon}/4)\cdot(\|f(x')-f(y')\| - 2L(R_x+R_y)) + 1\cdot(c\sqrt{\varepsilon}L(R_x+R_y))}{c\sqrt{\varepsilon}/4+1} \\ &= \frac{c\sqrt{\varepsilon}}{4+c\sqrt{\varepsilon}}(\|f(x')-f(y')\| + 2L(R_x+R_y)) \ge \frac{c\sqrt{\varepsilon}}{4+c\sqrt{\varepsilon}} d_{\mathrm{ub}}(x,y). \end{aligned}$$

4 Proof of Theorem 5

In this section, we prove Theorem 5. To prove item 1, consider 4 points in the plane: $x=(0,0)$, $y=(0,\sqrt{2})$, $u=(1,0)$, and $v=(1,\sqrt{2})$. Let f be the map that sends x,y,u, and v to $x'=(0,0)$, $y'=(1,1)$, $u'=(1,0)$, and $v'=(0,1)$. It is easy to see that f is 1-Lipschitz. Now let $z_1=(x+y)/2$ and $z_2=(u+v)/2$. Note that $\|x-z_1\| = \|y-z_1\| = \|u-z_2\| = \|v-z_2\| = \sqrt{2}/2$ and $d_{\mathrm{ub}}(z_1,z_2)=1$.

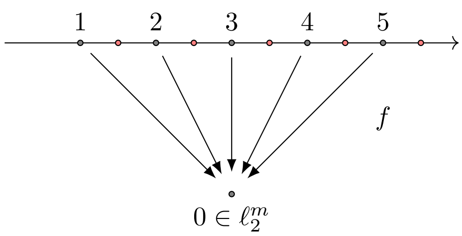


Consider a $(1+\varepsilon)$-Lipschitz proper or outer extension $\tilde{f}$ of f to $\{x,y,u,v,z_1,z_2\}$. Let $z_1' = \tilde{f}(z_1)$ and $z_2' = \tilde{f}(z_2)$. Since $\tilde{f}$ is $(1+\varepsilon)$-Lipschitz, we have $\|x'-z_1'\|^2 \le (1+\varepsilon)^2/2$ and $\|y'-z_1'\|^2 \le (1+\varepsilon)^2/2$. Hence

$$\left\|\frac{x'+y'}{2} - z_1'\right\|^2 = \frac{\|x'-z_1'\|^2 + \|y'-z_1'\|^2}{2} - \frac{\|x'-y'\|^2}{4} \le \frac{(1+\varepsilon)^2/2 + (1+\varepsilon)^2/2}{2} - \frac{1}{2} = \varepsilon + \frac{\varepsilon^2}{2}$$

Therefore, $\left\|\frac{x'+y'}{2} - z_1'\right\| = O(\sqrt{\varepsilon})$. Similarly, $\left\|\frac{u'+v'}{2} - z_2'\right\| = O(\sqrt{\varepsilon})$. Since $\frac{x'+y'}{2} = \frac{u'+v'}{2}$, we have $\|z_1'-z_2'\| = O(\sqrt{\varepsilon}) = O(\sqrt{\varepsilon}\, d_{\mathrm{ub}}(z_1',z_2'))$, as required.

Now we prove item 2. Consider sets $S=\{\kappa e_1 : 1\le\kappa\le N\}$ and $S'=\{(\kappa+{}^1\!/\!_2)e_1 : 1\le \kappa\le N\}$ in ℓ_2^n (where e_1 is the first standard basis vector in ℓ_2^n). Let $T = S\cup S'$.



Consider map f that sends S to 0 in ℓ_2^m. Trivially, map f is 1-Lipschitz. We have, $d_{\mathrm{ub}}(x,y) = \min(\|x-y\|, {}^1\!/_2 + 0 + {}^1\!/_2) = 1$ for every distinct $x, y \in S'$. Consider an L-Lipschitz extension $\tilde{f} : T \to \ell_2^{m'}$ such that $\|\tilde{f}(x) - \tilde{f}(y)\| \geq c d_{\mathrm{ub}}(x,y)$. Since every point x in S' is at distance $1/2$ from some point in S, $\tilde{f}$ maps all points in S' to a ball of radius ${}^L\!/_2$ around 0 in $\ell_2^{m'}$. The distance between the images of every two points is at least $c d_{\mathrm{ub}}(x,y) = c$. Therefore, $\tilde{f}(S')$ is a c-separated set in a ball of radius ${}^L\!/_2$ in $\ell_2^{m'}$. We apply a standard argument to bound the size of $f(S')$. Ball of radius $c/2$ around points in $\tilde{f}(S')$ are mutually disjoint and lie in a ball of radius ${}^{(L+c)}\!/_2$. Thus, the number of points in $\tilde{f}(S')$ by the ratio of the volumes of balls of radius ${}^{(L+c)}\!/_2$ and $c/2$. That is,

$$N = |\tilde{f}(S')| \leq \left(\frac{L+c}{c}\right)^{m'}.$$

It follows that $m' \geq \frac{\log_e N}{\log_e (L/c+1)}$, as required.

Finally, observe that item 3 follows from item 2 if we let $N = \infty$.

5 Proof of Theorem 7

In this section, we prove Theorem 7.

Proof of part I. We first assume that $C = 1$ and then consider the general case when $C \geq 1$ is arbitrary. Let g and h be embeddings of (X, d_X) and (Y, d_Y) into Euclidean spaces ℓ_2^n and ℓ_2^m with distortions at most D_X and D_Y, respectively. Without loss of generality, we may assume that g and h are 1-Lipschitz and that they do not contract distances by more than D_X and D_Y, respectively. Define map h' as $h'(y) = {}^{h(y)}\!/_{D_X}$. Let $S = g(Y)$ and $T = g(X)$. Consider map f that sends $u \in S$ to $h'(g^{-1}(u))$.

$$\begin{tikzcd}
\ell_2^n \supset S = g(Y) \arrow[d, "\subset"] \arrow[rr, bend left, "f = h' \circ g^{-1}"] & Y \arrow[l, "g"'] \arrow[r, "h' = \frac{h}{D_X}"] \arrow[d, "\subset"] & h'(Y) \subset \ell_2^m \arrow[d, "\subset"] \\
\ell_2^n \supset T = g(X) \arrow[rr, bend right, "\tilde{f}"'] & X \arrow[l, "g"'] \arrow[r, "\Phi"] & \ell_2^{m'}
\end{tikzcd}$$

Note that f is 1-Lipschitz, since

$$\begin{aligned}
\|f(u) - f(v)\| &= \|h'(g^{-1}(u)) - h'(g^{-1}(u))\| \leq \|h'\|_{Lip} \cdot d_Y(g^{-1}(u), g^{-1}(v)) \\
&\leq \frac{\|h\|_{Lip}}{D_X} \cdot d_X(g^{-1}(u), g^{-1}(v)) \leq \frac{1}{D_X} \cdot D_X \|u - v\| = \|u - v\|.
\end{aligned}$$

Consider $x, y \in X$. We show that $\frac{d_u(x,y)}{D_X D_Y} \leq d_{\mathrm{ub}}(g(x), g(y)) \leq d_u(x,y)$ where d_u is the updated metric on X, and d_{ub} is with respect to the map f. We have,

$$\begin{aligned}
d_{\mathrm{ub}}(g(x), g(y)) &= \min(\|g(x) - g(y)\|, \inf_{a,b \in Y} (\|g(x) - g(a)\| + \|h'(a) - h'(b)\| + \|g(b) - g(y)\|)) \\
&\leq \min(d_X(x,y), \inf_{a,b \in Y} (d_X(x,a) + \frac{d_Y(a,b)}{D_X} + d_X(b,y))) \\
&\leq \min(d_c(x,y), \inf_{a,b \in Y} (d_c(x,a) + d_c(a,b) + d_c(b,y))) = d_u(x,y)
\end{aligned}$$

where to get the last inequality, we need to do case analysis on the four cases given by whether $x \in Y$ or $x \notin Y$; and $y \in Y$ or $y \notin Y$. The last equality follows by considering the shortest path in d_c between x and y (of weight $d_u(x,y)$), and doing simple case analysis using the fact that $C = 1$. Then

$$\begin{aligned} d_{\mathrm{ub}}(g(x), g(y)) &= \min(\|g(x) - g(y)\|, \inf_{a,b \in Y}(\|g(x) - g(a)\| + \|h'(a) - h'(b)\| + \|g(b) - g(y)\|)) \\ &\geq \min\Big(\frac{d_X(x,y)}{D_X}, \inf_{a,b\in Y}\Big(\frac{d_X(x,a)}{D_X} + \frac{d_Y(a,b)}{D_Y D_X} + \frac{d_X(b,y)}{D_Y}\Big)\Big) \geq \frac{d_u(x,y)}{D_X D_Y}. \end{aligned}$$

By Theorem 4, there exists an $3/2$-Lipschitz extension $\tilde{f} : T \to \ell_2^{m'}$ of f such that $\|\tilde{f}(u) - \tilde{f}(v)\| = \Theta(d_{\mathrm{ub}}(u,v))$ for $u, v \in T$. Consider $\Phi = \tilde{f} \circ g$, which maps X to $\ell_2^{m'}$. We obtain

$$\Omega\left(\frac{d_u(x,y)}{D_X D_Y}\right) \leq \|\Phi(x) - \Phi(y)\| \leq O(d_u(x,y)).$$

We conclude that X equipped with the updated metric d_u embeds into $\ell_2^{m'}$ with distortion at most $O(D_X D_Y)$.

To get the result for an arbitrary $C \geq 1$, we consider metric d'_Y defined by $d'_Y(x,y) = d_Y(x,y)/C$. Note that d'_Y is also D_Y-Euclidean. Additionally, $d'_Y(x,y) \leq d_X(x,y)$ for every $x, y \in Y$. Thus the updated metric d'_u for d_X and d'_Y is $O(D_X D_Y)$-Lipschitz. Finally, we note that $d'_u(x,y) \leq d_u(x,y) \leq C d'_u(x,y)$ and hence metric d_u is $O(C\, D_X D_Y)$-Euclidean. ◀

Proof of part II. Let p be the largest prime number such that $2p \leq N$. We construct an expander $G = (V, E)$ on p vertices, which is a union of two Hamiltonian paths (the paths are not necessarily disjoint). We note that any such expander will work, but we will describe one to be more specific. Let $V = \mathbb{Z}/p\mathbb{Z}$, $E_1 = \{(i, i+1) : i \in \{0, \dots, p-2\}$, and $E'_2 = \{(i,j) : i \cdot j = 1,\ i \neq j \text{ where } i, j \in \mathbb{Z}/p\mathbb{Z}\}$ (where the product $i \cdot j$ is computed in $\mathbb{Z}/p\mathbb{Z}$). Observe that E'_2 is a partial matching, so we can choose a set of edges E''_2 such that $E'_2 \cup E''_2$ is a path visiting all vertices in V once. Now, $G = (V, E_1 \cup E'_2 \cup E''_2)$; P_1 and P_2 are Hamiltonian paths with edge sets E_1 and $E_2 = E'_2 \cup E''_2$, respectively. Graph $(V, E_1 \cup E'_2)$ is an expander (see [13] or Construction 4.26 in [18]) and thus so is G. Denote the shortest path distance in G by d_G.

Now, for every $i \in V$, we create two points u_i and u'_i. Let $X = \{u_i, u'_i : i \in V\}$ and $Y = \{u'_i : i \in V\}$. Consider path P_2 and one of its endpoints. Let $\pi(i)$ be the distance from vertex u_i to this endpoint along P_2. Let $\varepsilon = 1/p$. Now we define distances d_X and d_Y:

$$\begin{aligned} &d_X(u_i, u_j) = d_X(u'_i, u'_j) = |i - j| \text{ and } d_X(u_i, u'_j) = \sqrt{|i-j|^2 + \varepsilon^2}, \\ &d_Y(u'_i, u'_j) = |\pi(i) - \pi(j)|. \end{aligned}$$

Here, the value of $|i - j|$ is computed in $\mathbb{Z}$, not in $\mathbb{Z}/p\mathbb{Z}$. Observe that (X, d_X) and (Y, d_Y) embed into ℓ_2^2 and $\ell_2^1 = \mathbb{R}$ isometrically. Indeed, the map that sends u_i to $(i, 0)$ and u'_i to (i, ε) is an isometric embedding of X into the Euclidean plane; the map that sends u'_i to $\pi(i) \in \mathbb{R}$ is an isometric embedding of Y into the real line. Now consider the combined and updated metrics, d_c and d_u, on X. We have,

- $d_c(u_i, u_j) = d_X(u_i, u_j) = |i - j| \geq d_G(i,j)$;
- $d_c(u_i, u'_j) = d_X(u_i, u'_j) = \sqrt{|i-j|^2 + \varepsilon^2} > d_G(i,j)$
- $d_c(u'_i, u'_j) = d_Y(u'_i, u'_j) = |\pi(i) - \pi(j)| \geq d_G(i,j)$.

Let us derive lower and upper bounds on $d_u(u_i, u_j)$ in terms of $d_G(i, j)$. Fix some i and j. Denote the complete graph on X with edge lengths $d_c(\cdot, \cdot)$ by $K(X, d_c)$; denote the complete graph on V with edge lengths $d_G(\cdot, \cdot)$ by $K(V, d_G)$. Note that the shortest path distance in $K(V, d_G)$ equals d_G.

By definition, $d_u(u_i, u_j)$ is the shortest path distance between u_i and u_j in $K(X, d_c)$. Consider a shortest path between u_i and u_j in $K(X, d_c)$ and its "projection" to $K(V, d_G)$, which we obtain as follows. We replace each vertex on the path with a corresponding vertex in V: namely, we replace u_a and u'_a with a. We get a path from i to j in $K(V, d_G)$; its length is at least $d_G(i, j)$. Since $d_c(u_a, u_b)$, $d_c(u'_a, u_b)$, and $d_c(u'_a, u'_b)$ are all greater than or equal to $d_G(a, b)$, the length of the projected path in $K(V, d_G)$ is at most the length of the path between u_i and u_j in $K(X, d_c)$. Therefore, the shortest path distance between u_i and u_j in $K(X, d_c)$ is at least $d_G(i, j)$. We conclude that $d_u(u_i, u_j) \geq d_G(i, j)$.

Now consider a shortest path P between i and j in G. We "lift" it to $K(X, d_c)$ as follows. We replace each edge $(a, b) \in E_1$ of P with edge (u_a, u_b) in $K(X, d_c)$ and each edge $(a, b) \in E_2 \setminus E_1$ with edge (u'_a, u'_b). Note that this transformation preserves edge lengths; all of these edges have length 1. We obtain a sequence of edges in $K(X, d_c)$, which does not necessarily form a path, since it may happen that one edge ends at u_a and the next one starts at u'_a (or vice versa). We transform it to a path between u_i and u_j in $K(X, d_c)$ by adding edges of the form (u_a, u'_a) where necessary. The length of the lifted path equals the length of P, which is $d_G(i, j)$, plus the length of all the additional edges (u_a, u'_a). Each of the additional edges has length ε and there are at most p of them. We conclude that there is a path of length at most $d_G(i, j) + 1$ between u_i and u_j in $K(X, d_c)$. Thus $d_u(u_i, u_j) \leq d_G(i, j) + 1$.

We proved that $d_u(u_i, u_j) = \Theta(d_G(i, j))$. Because G is an expander, every embedding of (V, d_G) into Euclidean space requires distortion $\Omega(\log N)$ [11]. Therefore, every embedding of $(X \setminus Y, d_u)$ and, consequently, of (X, d_u) into ℓ_2 also requires distortion at least $\Omega(\log N)$. ◀

References

1 Mihai Badoiu, Kedar Dhamdhere, Anupam Gupta, Yuri Rabinovich, Harald Räcke, Ramamoorthi Ravi, and Anastasios Sidiropoulos. Approximation algorithms for low-distortion embeddings into low-dimensional spaces. In *Proceedings of the Symposium on Discrete Algorithms*, volume 5, pages 119–128, 2005.

2 Yair Bartal, Ben Recht, and Leonard J Schulman. Dimensionality reduction: beyond the Johnson–Lindenstrauss bound. In *Proceedings of the Symposium on Discrete Algorithms*, pages 868–887. SIAM, 2011.

3 Armin Biess, Aryeh Kontorovich, Yury Makarychev, and Hanan Zaichyk. Regression via Kirszbraun extension with applications to imitation learning. *CoRR*, abs/1905.11930, 2019. arXiv:1905.11930.

4 Yair Censor, Aviv Gibali, and Simeon Reich. Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. *Optimization Methods and Software*, 26(4-5):827–845, 2011.

5 Nishant Chandgotia, Igor Pak, and Martin Tassy. Kirszbraun-type theorems for graphs. *Journal of Combinatorial Theory, Series B*, 137:10–24, 2019.

6 Timothy Chu, Gary L Miller, and Donald R Sheehy. Exact computation of a manifold metric, via lipschitz embeddings and shortest paths on a graph. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 411–425. SIAM, 2020.

7 Lee-Ad Gottlieb, Aryeh Kontorovich, and Robert Krauthgamer. Efficient regression in metric spaces via approximate Lipschitz extension. *IEEE Transactions on Information Theory*, 63(8):4838–4849, 2017.

8 William Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984.

9 Mojzesz D. Kirszbraun. Über die zusammenziehende und Lipschitzsche Transformationen. *Fundamenta Mathematicae*, 22:77–108, 1934.

10 James R Lee and Assaf Naor. Extending Lipschitz functions via random metric partitions. *Inventiones mathematicae*, 160(1):59–95, 2005.

11 Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

12 Steven H Low. Analytical methods for network congestion control. *Synthesis Lectures on Communication Networks*, 10(1):1–213, 2017.

13 Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

14 Sepideh Mahabadi, Konstantin Makarychev, Yury Makarychev, and Ilya Razenshteyn. Nonlinear dimension reduction via outer bi-Lipschitz extensions. In *Proceedings of the Symposium on Theory of Computing*, pages 1088–1101, 2018.

15 Konstantin Makarychev and Yury Makarychev. A union of Euclidean metric spaces is Euclidean. *Discrete Analysis*, 14, 2016.

16 Konstantin Makarychev, Yury Makarychev, and Ilya Razenshteyn. Performance of Johnson–Lindenstrauss transform for k-means and k-medians clustering. In *Proceedings of the Symposium on Theory of Computing*, pages 1027–1038, 2019.

17 Manor Mendel and Assaf Naor. Euclidean quotients of finite metric spaces. *Advances in Mathematics*, 189(2):451–494, 2004.

18 Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1–3):1–336, 2012.