Fréchet embeddings of negative type metrics

Sanjeev Arora* James R. Lee† Assaf Naor ‡

Abstract

We show that every n-point metric of negative type (in particular, every n-point subset of L_1) admits a Fréchet embedding into Euclidean space with distortion $O\left(\sqrt{\log n} \cdot \log \log n\right)$, a result which is tight up to the $O(\log \log n)$ factor, even for Euclidean metrics. This strengthens our recent work on the Euclidean distortion of metrics of negative into Euclidean space.

1 Introduction

Let (X, d_X) and (Y, d_Y) be finite metric spaces. Given an injection $f : X \hookrightarrow Y$, the *distortion* of f is defined as:

$$\mathrm{dist}(f) \coloneqq \|f\|_{\mathrm{Lip}} \cdot \|f^{-1}\|_{\mathrm{Lip}} = \sup_{\substack{x,y \in X \\ x \neq Y}} \frac{d_Y(f(x), f(y))}{d_X(x,y)} \cdot \sup_{\substack{x,y \in X \\ x \neq Y}} \frac{d_X(x,y)}{d_Y(f(x), f(y))}.$$

The least distortion with which X may be embedded into Y is denoted by $c_Y(X)$, i.e.

$$c_Y(X) \coloneqq \inf\{\mathrm{dist}(f) : \ f : X \hookrightarrow Y\}.$$

For $p \geq 1$ we also use the notation $c_p(X) \coloneqq c_{L_p}(X)$. The parameter $c_2(X)$ is called the *Euclidean distortion* of X.

Bourgain's fundamental embedding theorem [6] states that:

$$|X| = n \implies c_2(X) = O(\log n). \tag{1}$$

But, Bourgain's proof of (1) contains more information. A *Fréchet embedding* of (X, d_X) is a probability distribution μ over all non-empty subsets of X. If A is a random subset distributed according to μ, then we associate, to every $x \in X$, the real-valued random variable $F_\mu(x) = d(x, A)$.

Thus, for every $x, y \in X$ and every $p > 0$,

$$\|F_\mu(x) - F_\mu(y)\|_p = \left(\mathbb{E}_\mu \left|d_X(x, A) - d_X(y, A)\right|^p\right)^{1/p}.$$

In [6] Bourgain constructs a Fréchet embedding μ such that for every $x, y \in X$,

$$\|F_\mu(x) - F_\mu(y)\|_\infty \leq d_X(x,y) \leq O(\log n) \cdot \|F_\mu(x) - F_\mu(y)\|_1. \tag{2}$$

*Supported by David and Lucile Packard Fellowship and NSF grant CCR-0205594. Princeton University. arora@cs.princeton.edu

†Supported by NSF grant CCR-0121555 and an NSF Graduate Research Fellowship. U. C. Berkeley. jrl@cs.berkeley.edu

‡Microsoft Research. anaor@microsoft.com

Observe that (2) implies (1), and that the left-hand inequality in (2) holds automatically, since by the triangle inequality, the mapping F_μ is *pointwise* 1-Lipschitz. This fundamental property of Bourgain's embedding is crucial for certain applications—for example it is used in the design of approximation algorithms for vertex separators [9]. From an analytic point of view, the "mixed norm" inequality (2) is natural since it can be viewed as a (non-linear) Dvoretzky-Rogers type embedding (see [8, 10]).

Unfortunately, it isn't always possible to construct Fréchet embeddings. For example, in [16] it is shown that if $\mathcal{N}$ is a $1/\sqrt{d}$ net on the unit d-dimensional Euclidean sphere, then for every Fréchet embedding $F_\mu : \mathcal{N} \to L_2(\mu)$,

$$\mathrm{dist}(F_\mu) = \Omega\left(\sqrt{d}\right) = \Omega\left(\sqrt{\frac{\log|\mathcal{N}|}{\log\log|\mathcal{N}|}}\right).$$

Thus, we cannot expect to have a Fréchet embedding for any n-point Euclidean metric with Euclidean distortion significantly better than $\sqrt{\log n}$. In the context of metric Ramsey (non-linear Dvoretzky) embeddings [5], it was shown in [4] that Fréchet embeddings cannot be used to obtain the results of [5] (we refer to these papers for more details). In [17] it was shown that any n-point weighted planar graph (equipped with the shortest-path metric) embeds into L_2 with distortion $O\left(\sqrt{\log n}\right)$. The embedding of [17] is not a Fréchet embedding, and it required significantly more work to prove in [13] that Fréchet embeddings exist with the same distortion guarantee.

Recall that (X, d_X) is said to be a metric space of *negative type* if the metric space $(X, \sqrt{d_X})$ embeds isometrically into Euclidean space. The space L_1 has negative type, and metrics of negative type also occur as relaxations of certain semidefinite programs (see e.g. [11]). It is known [12] that there are metrics of negative type that require arbitrarily large distortion in any embedding into L_1. In [1] the present authors proved that if (X, d_X) is an n-point metric space of negative type then $c_2(X) = O\left(\sqrt{\log n} \cdot \log\log n\right)$. This result is optimal up to the iterated logarithm, and is used in [1] to obtain the best know approximation algorithm for the Sparsest Cut problem with general demands (see the discussion in [1] for background on this topic). On the other hand, the best known bound for Fréchet embeddings of negative type metrics is $O(\log n)^{3/4}$ [7].

The embedding of [1] is *not* a Fréchet embedding. The argument of [1] is modular and general, as it presents a gluing technique for certain ensembles of Lipschitz mappings. The resulting "glued" mapping is not Fréchet, even if the original ensemble consists of Fréchet embeddings. While this general gluing procedure is of independent interest, it is natural to ask whether it is also possible to obtain a Fréchet embedding with the same distortion guarantee. In this paper we show that this is indeed the case, by proving the following theorem

Theorem 1.1. *Let (X, d_X) be an n-point metric space of negative type. Then there exists a probability measure μ over random subsets $\emptyset \neq A \subseteq X$ such that for every $x, y \in X$,*

$$\left(\mathbb{E}_\mu |d(x,A) - d(y,A)|^2\right)^{1/2} = \Omega\left(\frac{d_X(x,y)}{\sqrt{\log n} \cdot \log\log n}\right).$$

We remark that this bound is new even for the special case when X is an n-point submetric of some Euclidean space. The proof of Theorem 1.1 is different, and substantially more involved than the proof in [1]. We believe that it is worthwhile to establish that Fréchet embeddings are achievable in this case. Such maps are interesting due to the algorithmic and combinatorial applications of

Fréchet embeddings in [9], in addition to the new structural information contained in Theorem 1.1. Moreover, the proof techniques used here are different than in [1], and are independently interesting.

We end this introduction by stating some interesting open problems related to Fréchet embeddings. Our result suggests (together with the results of [13]) that Fréchet embeddings may be universal for large enough distortions. More precisely, we have the following question.

Question 1. Does *every* n-point metric space (X, d_X) admits a Fréchet embedding into L_2 with distortion $O\left(\max\{c_2(X), \sqrt{\log n}\}\right)$?

Observe that $\sqrt{\log n}$ is a natural barrier here due to the result of [16]. If this (speculative) bound is indeed true, then it would yield a method of producing good Fréchet embeddings from general embeddings (and it would show that Theorem 1.1 follows from [1]).

Finally, it is not known whether every n-point metric of negative type admits a Fréchet embedding as in the statement of Theorem 1.1 with the 2-norm replaced by the 1-norm.

Question 2. Does every n-point metric space of negative type admit a Fréchet embedding into L_1 with distortion $O(\sqrt{\log n} \log \log n)$?

If true, this would yield the best known approximation algorithm for the minimum-weight vertex separator problem (see [9]). In fact, the best upper bound known to hold is Bourgain's bound of $O(\log n)$. This question is open even for the shortest-path metrics of planar graphs, and its resolution has implications for the theory of vertex-capacitated flows in such families.

2 Preliminaries

We recall the following definition from [1]. Let (X, d) be an n-point metric space.

Definition 2.1 (Random zero-sets). *Given $\Delta, \zeta > 0$, and $p \in (0,1)$ we say that X admits a random zero set at scale Δ which is ζ-spreading with probability p if there is a distribution μ over subsets $Z \subseteq X$ such that for every $x, y \in X$ with $d(x,y) \geq \Delta$,*

$$\mu\left\{Z \subseteq X :\ y \in Z \text{ and } d(x, Z) \geq \frac{\Delta}{\zeta}\right\} \geq p.$$

We denote by $\zeta(X;p)$ the least $\zeta > 0$ such that for every $\Delta > 0$, X admits a random zero set at scale Δ which is ζ-spreading with probability p. Finally, given $k \leq n$ we define

$$\zeta_k(X;p) = \max_{\substack{Y \subseteq X \\ |Y| \leq k}} \zeta(Y;p).$$

As noted in [1], a concatenation of the results of [2], [14], and [7], shows that there exists a universal constant $p \in (0,1)$ such that for every n-point metric space (X,d) of negative type, $\zeta(X;p) = O\left(\sqrt{\log n}\right)$.

We now recall the related notion of *padded decomposability.* Given a partition P of X and $x \in X$ we denote by $P(x) \in P$ the unique element of P to which x belongs. In what follows we sometimes refer to $P(x)$ as the *cluster* of x. Following [13] we define the *modulus of padded decomposability* of X, denoted α_X, as the least constant $\alpha > 0$ such that for every $\Delta > 0$ there is a distribution ν over partitions of X with the following properties.

1. For all $P \in \mathrm{supp}(\nu)$ and all $C \in P$ we have that $\mathrm{diam}(C) < \Delta$.

2. For every $x \in X$ we have that

$$\nu\{P : \ B(x, \Delta/\alpha) \subseteq P(x)\} \geq \frac{1}{2}.$$

As observed in [13], the results of [15, 3] imply that $\alpha_X = O(\log |X|)$, and this will be used in our proof. Moreover, it is shown in [1] (motivated by an argument in [17]) that it is always the case that $\zeta(X; 1/8) \leq \alpha_X$.

We conclude this section with the following elementary probabilistic lemma.

Lemma 2.2 (Sampling lemma). *Suppose that $k, n \in \mathbb{N}$, $1 \leq k \leq n$. Let X be an n-point set and let $\widetilde{X}_k$ be chosen uniformly at random from all k-point subsets of X. Then*

1. *For every $A \subseteq X$, $\Pr[\widetilde{X}_k \cap A = \emptyset] \leq e^{-k|A|/n}$.*

2. *For every $A, B \subseteq X$ such that $A \cap B = \emptyset$,*

   (a) $\Pr[\widetilde{X}_k \cap A = \emptyset \,|\, \widetilde{X}_k \cap B \neq \emptyset] \leq \Pr[\widetilde{X}_{k-1} \cap A = \emptyset] \leq e^{(1-k)|A|/n}$

   (b) $\Pr[\widetilde{X}_k \cap A = \emptyset \,|\, \widetilde{X}_k \cap B = \emptyset] \leq e^{-k|A|/n}$.

Proof. The proof of (1) is an easy calculation. To prove 2(a), note that choosing $\widetilde{X}_k$ uniformly subject to $\widetilde{X}_k \cap B \neq \emptyset$ is the same as first choosing $z \in B$ uniformly at random, then choosing a uniform $k-1$ point subset $S \subseteq X \setminus \{z\}$ and returning $S \cup \{z\}$. □

3 Proof of Theorem 1.1

The main technical result of this paper is contained in the following lemma.

Lemma 3.1 (Enhanced descent). *Let (X, d) be an n-point metric space and fix $p \leq 1/8, K \geq 2$ and $\zeta \geq \zeta_K(X; p)$. For every $m \in \mathbb{Z}$, define*

$$S_m(K) = \left\{ x \in X : \left| B\left(x, 2^{m+5}\alpha_X\right) \right| \leq \frac{K}{16} \cdot \left| B\left(x, \frac{2^{m-9}}{\zeta}\right) \right| \right\}.$$

Then there exists a distribution σ over random subsets $A \subseteq X$ such that for all $m \in \mathbb{Z}$, $x \in S_m(K)$ and $y \in X$ with $d(x, y) \in [2^{m-1}, 2^m]$,

$$\mathbb{E}_\sigma |d(x, A) - d(y, A)|^2 \geq \frac{p^5}{O(\log n \log\log n)} \cdot \frac{d(x, y)^2}{\zeta^2} \cdot \log \frac{|B(x, 2^{m+5}\alpha_X)|}{|B(x, 2^{m+3}/\zeta)|}. \tag{3}$$

Before proving Lemma 3.1 we show how it implies Theorem 1.1. The argument below actually yields more general results. For example if we assume that $\zeta_k(X; p) = O(\log k)^\theta$ for some $p \in (0, 1/8)$, $\theta \geq \frac{1}{2}$ and all $k \leq n$ then we achieve a Fréchet embedding which achieves

$$c_2(X) = O_p\left((\log n)^\theta \sqrt{\log \alpha_X \log\log n} \right) = O_p\left((\log n)^\theta \log\log n \right),$$

where $O_p(\cdot)$ may contain an implicit constant which depends only on p.

Proof of Lemma 3.1 $\Longrightarrow$ Theorem 1.1. As noted above, there exists a universal constant $p \in (0,1)$ such that for every metric space (Y,d) of negative type, $\zeta(Y;p) = O\left(\sqrt{\log |Y|}\right)$. Combining this with Lemma 3.1 and the fact that $\alpha_X = O(\log n)$ we obtain the following statement. There exists a constant $C > 0$ such that for every $K \geq 2$ there exists a distribution μ_K over random subsets $A_K \subseteq X$ satisfying the following condition. Define

$$S'_m(K) = \left\{ u \in X : \left|B\left(u, 2^{m+5}\alpha_X\right)\right| \leq K \cdot \left|B\left(u, \frac{2^m}{C\sqrt{\log K}}\right)\right| \right\}.$$

Then for all $m \in \mathbb{Z}$, $x \in S'_m(K)$ and $y \in X$ such that $d(x,y) \in [2^{m-1}, 2^m]$,

$$\mathbb{E}_{\mu_K} |d(x,A_K) - d(y,A_K)|^2 \geq \frac{d(x,y)^2}{C \log K \log n \log\log n} \cdot \log \frac{|B(x, 2^{m+5}\alpha_X)|}{|B(x, C2^m/\sqrt{\log K})|}.$$

Observe that for every $m \in \mathbb{Z}$, $S'_m(n) = X$. Hence, defining $K_0 = n$ and $K_{j+1} = K_j^{1/C^4}$, as long as $K_j \geq 2$, we obtain random subsets $U_0, U_1, \ldots, U_j \subseteq X$ (where $U_j = A_{K_j}$) satisfying, for all $x \in S_m(K_j) \setminus S_m(K_{j+1})$ and $y \in X$ such that $d(x,y) \in [2^{m-1}, 2^m]$,

$$\begin{aligned}
\mathbb{E}\,|d(x,U_j) - d(y,U_j)|^2 &\geq \frac{d(x,y)^2}{C \log K_j \log n \log\log n} \cdot \log \frac{|B(x, 2^{m+5}\alpha_X)|}{|B(x, C2^m/\sqrt{\log K_j})|} \\
&= \frac{d(x,y)^2}{C \log K_j \log n \log\log n} \cdot \log \frac{|B(x, 2^{m+5}\alpha_X)|}{\left|B\left(x, C2^m/\sqrt{\log K_{j+1}^{C^4}}\right)\right|} \\
&= \frac{d(x,y)^2}{C \log K_j \log n \log\log n} \cdot \log \frac{|B(x, 2^{m+5}\alpha_X)|}{|B(x, 2^m/(C\sqrt{\log K_{j+1}}))|} \\
&\geq \frac{d(x,y)^2}{C \log K_j \log n \log\log n} \cdot \log K_{j+1} = \frac{d(x,y)^2}{C^5 \log n \log\log n}.
\end{aligned}$$

This procedure ends after N steps, where $N \leq \frac{\log\log n}{\log C}$. Every $x \in S_m(K_N)$ satisfies

$$|B(x, 2^{m+5}\alpha_X)| \leq 2^{C^4} |B(x, 2^{m+1}/C)|.$$

By a result of [13] (for a simple proof, see [1, Claim 4.6]) there is a distribution over random subsets $U_{N+1} \subseteq X$ such that for every $x, y \in S_m(K_N)$,

$$\mathbb{E}\,|d(x,U_{N+1}) - d(y,U_{N+1})|^2 \geq \frac{d(x,y)^2}{O(\log n)}.$$

Now consider the distribution μ over random subsets $U \subseteq X$, defined as follows. First, pick a value $j \in \{0,1,\ldots,N+1\}$ uniformly at random, then choose a random subset from the distribution of U_j. For every $x, y \in X$ choose $m \in \mathbb{Z}$ such that $d(x,y) \in [2^{m-1}, 2^m]$. If $x, y \in S_m(K_N)$ then

$$\mathbb{E}_\mu |d(x,U) - d(y,U)|^2 \geq \frac{\mathbb{E}\,|d(x,U_{N+1}) - d(y,U_{N+1})|^2}{N+1} \geq \frac{d(x,y)^2}{O(N \log n)}.$$

Otherwise, there is $j \in \{0, \ldots, N-1\}$ such that $x \in S_m(K_j) \setminus S_m(K_{j+1})$, in which case

$$\mathbb{E}_\mu |d(x,U) - d(y,U)|^2 \geq \frac{\mathbb{E}\,|d(x,U_{j+1}) - d(y,U_{j+1})|^2}{N+1} \geq \frac{d(x,y)^2}{O(N \log n \log\log n)}.$$

Recalling that $N = O(\log\log n)$ completes the proof. $\square$

3.1 Proof of Lemma 3.1: Enhanced descent

We begin with a simple definition.

Definition 3.2. *For every $x \in X$ and $t > 0$ define*

$$\kappa(x,t) = \max\{\kappa \in \mathbb{Z} : |B(x, 2^{\kappa})| < 2^t\}. \tag{4}$$

The following simple lemma says that the values of $\kappa(\cdot,\cdot)$ cannot change too rapidly when moving between nearby points. This fact will be used several times in the ensuing arguments, and played a similar role in [13].

Lemma 3.3 (Smoothness). *For $x \in X$, let $i \in \mathbb{Z}$ and $m, t \in \mathbb{Z}^+$ be such that $\left|B(x, 2^{i+m-1})\right| \le 2^t \le \left|B(x, 2^{i+m})\right|$. Then every $z \in X$ for which $d(x,z) \le 2^{\min\{m, m+i-2\}}$ satisfies:*

$$\kappa(z,t) \in \{m+i-3, m+i-2, m+i-1, m+i, m+i+1\}.$$

Proof. By definition,

$$|B(z, 2^{\kappa(z,t)})| < 2^t \le |B(x, 2^{\kappa(z,t)+1})|.$$

For the upper bound, we have

$$|B(x, 2^{\kappa(z,t)} - 2^{m-s})| \le |B(z, 2^{\kappa(z,t)})| < 2^t \le |B(x, 2^{i+m})|,$$

implying that $2^{\kappa(z,t)} - 2^{m-s} < 2^{i+m}$, which yields $2^{\kappa(z,t)} < 2^{m+1+i}$. For the lower bound, we have

$$|B(x, 2^{\kappa(z,t)+1} + 2^{m-s})| \ge |B(z, 2^{\kappa(z,t)+1})| \ge 2^t \ge |B(x, 2^{m+i-1})|.$$

We conclude that $2^{\kappa(z,t)+1} + 2^{m-s} \ge 2^{m+i-1}$, which implies that $2^{\kappa(z,t)+1} \ge 2^{m+i-2}$. □

Notation. We introduce some notation which will be used in the forthcoming proofs. Write $\alpha = \alpha_X$, and for every $j \in \mathbb{Z}$ let P_j denote a random partition of X satisfying the following.

1. For all $C \in P_j$ we have that $\mathrm{diam}(C) \le 2^{j+4}\alpha$.
2. For every $x \in X$ we have that $\nu\{P : \ B(x, 2^{j+4}) \subseteq P_j(x)\} \ge \frac{1}{2}$.

We also fix $p \in (0, 1/8)$ and for every $k \le n$ let $\zeta_k = \zeta_k(X; p)$. For $S \subseteq X$ let $\Psi_j(S)$ denote a random zero set of S at scale 2^{j-3} which is $\zeta_{|S|}$-spreading with probability p. For each $C \subseteq X$ let $\widetilde{C}$ be a uniformly random subset of C of size $\min\{K, |C|\}$.

The Fréchet-type embeddings. The embeddings we produce will be of Fréchet-type, i.e. every coordinate $f_i : X \to \mathbb{R}$ will be of the form $f_i(x) = d(x, U)$ for some $U \subseteq X$. Let $I = [-\log_2 \zeta_K + 3, \log_2 \alpha + 6] \cap \mathbb{Z}$ and $J = \{0, 1, \ldots, \lceil \log_2 n \rceil\}$. For each $i \in I$ and $t \in J$, we describe a distribution W_t^i on sets. Our random embedding consists of mapping x to $f(x) = (d(x, W_t^i) : i \in I, t \in J)$. Such a mapping is clearly Lipschitz with constant $\sqrt{|I| \cdot |J|} = O(\sqrt{\log n \cdot \log \alpha})$ (here we use the fact that $\zeta_K \le \alpha$).

Let $\{\sigma_m\}_{m \in \mathbb{Z}}$ be a sequence of random variables taking each of the values $\{0, 1, 2\}$ with probability $\frac{1}{3}$, which are independent of all the other random variables appearing in this proof. (In general, the reader should assume that samplings from various distributions are independent of one another.) We define the random subset W_t^i as follows.

$$x \in W_t^i \iff \begin{cases} x \in X & \text{if } \sigma_{\kappa(x,t)-i} = 0, \\ x \in \Psi_{\kappa(x,t)-i}\left(\widetilde{P_{\kappa(x,t)-i}(x)}\right) & \text{if } \sigma_{\kappa(x,t)-i} = 1, \\ x \in \widetilde{P_{\kappa(x,t)-i}}(x) & \text{if } \sigma_{\kappa(x,t)-i} = 2. \end{cases}$$

For the rest of the proof, let m be any integer, fix $x, y \in X$ such that $d(x,y) \in [2^{m-1}, 2^m]$, and assume that $x \in S_m(K)$. Let $s_i = \log_2 |B(x, 2^{i+m})|$, with $s_{\min I}$ and $s_{\max I}$ corresponding to the smallest and largest $i \in I$. The next claim follows from the "smoothness" of Lemma 3.3.

Claim 3.4. *For $i \in I$ and all $t \in \mathbb{Z} \cap [s_{i-1}, s_i]$, every $w \in B(x, 2^m/\zeta_K)$ satisfies*

$$m - 3 \leq \kappa(w,t) - i \leq m + 1.$$

Now we define

$$N(x) \coloneqq |\{(i,t) : i \in I, t \in [s_{i-1}, s_i] \cap \mathbb{Z}\}| = |\{t : t \in [s_{\min I}, s_{\max I}] \cap \mathbb{Z}\}|.$$

Observe that

$$N(x) \geq \log_2 \frac{|B(x, 2^{m+5}\alpha)|}{|B(x, 2^{m+3}/\zeta_K)|}.$$

We are going to get a contribution to $||f(x) - f(y)||_2^2$ from the sets $\{W_t^i\}$ where $t \in \mathbb{Z} \cap [s_{i-1}, s_i]$ for some $i \in I$. The number of such pairs is $N(x)$. Thus clearly we get the desired lower bound (3) if we can prove that for these values of i and t, we have

$$\mathbb{E}\,|d(x, W_t^i) - d(y, W_t^i)|^2 \geq \left(\frac{p}{64}\right)^5 \cdot \frac{2^{2m}}{\zeta_K}. \tag{5}$$

To prove (5) we fix $i \in I, t \in [s_{i-1}, s_i] \cap \mathbb{Z}$ and let

$$M = \{m-3, m-2, m-1, m, m+1\}$$

be the range of values from Claim 3.4.

3.1.1 Partitions and padding

For any $j \in M$ we have that $\operatorname{diam}(P_j(x)) \leq 2^{j+4}\alpha \leq 2^{m+5}\alpha$, so $B(x, 2^{m+5}\alpha) \supseteq P_j(x)$. Since $x \in S_m(K)$, it follows that $|P_j(x)| \leq \frac{K}{16}|B(x, 2^{m-9}/\zeta_K)|$. Recall that for $j \in M$ the random partition P_j satisfies

$$\Pr[d(x, X \setminus P_j(x)) \geq 2^{m+1}] \geq \Pr[d(x, X \setminus P_j(x) \geq 2^{j+4}] \geq \tfrac{1}{2}.$$

Define the event

$$\mathcal{E}_{\text{pad}}^j = \{d(x, X \setminus P_j(x)) \geq 2^{m+1}\},$$

and let

$$\mathcal{E}_{\text{pad}} = \bigcap_{j \in M} \mathcal{E}_{\text{pad}}^j.$$

Note that, by independence, we have $\Pr[\mathcal{E}_{\text{pad}}] \geq 2^{-5}$.

Suppose that $\mathcal{E}^j_{\text{pad}}$ occurs, then since $d(x,y) \leq 2^m$, we have $y \in P_j(x)$, implying $P_j(x) = P_y(x)$. Furthermore, since

$$B(x, 2^{m-9}/\zeta_K) \subseteq P_j(x),$$

when we sample down $P_j(x)$ to $\widetilde{P_j(x)}$, a set of size at most K, Lemma 2.2, part (1), ensures that

$$\Pr[\widetilde{P_j(x)} \cap B(x, 2^{m-9}/\zeta_K) = \emptyset] \leq e^{-15}.$$

To this end, we denote

$$\mathcal{E}^j_{\text{hit}} = \left\{ \widetilde{P_j(x)} \cap B(x, 2^{m-9}/\zeta_K) \neq \emptyset \right\},$$

and we define $\mathcal{E}_{\text{hit}} = \bigcap_{j \in M} \mathcal{E}^j_{\text{hit}}$. Since the events $\{\mathcal{E}^j_{\text{hit}}\}_{j \in M}$ are independent even after conditioning on $\mathcal{E}_{\text{pad}}$, the preceding discussion yields the following lemma.

Lemma 3.5. $\Pr\left(\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}}\right) \geq 2^{-5}(1 - 5e^{-15}) > 2^{-6}$.

3.1.2 Obtaining a separation

We introduce the following events which mark different "phases" of the embedding. For $\ell \in \{1, 2\}$, let

$$\mathcal{E}^\sigma_\ell = \{\sigma_j = \ell \text{ for all } j \in M\}.$$

Note that $\Pr[\mathcal{E}^\sigma_\ell] \geq 3^{-5}$ for each $\ell \in \{1,2\}$. Now we study the distance from x to W^i_t in phase 1.

Claim 3.6. *If $\mathcal{E}^\sigma_1 \cap \mathcal{E}_{\text{pad}}$ occurs, then*

$$d(x, W^i_t) \geq \min\left\{ \frac{2^m}{\zeta_K}, \min_{j \in M} \left\{ d(x, \Psi_j(\widetilde{P_j(x)})) \right\} \right\}. \tag{6}$$

Proof. Fix a point $w \in B(x, 2^m/\zeta_K)$, and let $j = \kappa(w,t) - i$. By Claim 3.4, $j \in M$, hence $\mathcal{E}_{\text{pad}}$ implies that $w \in P_j(x)$. Since $\mathcal{E}^\sigma_1$ occurs, we have $w \in W^i_t$ if and only if $w \in \Psi_j(\widetilde{P_j(x)})$. □

If $\mathcal{E}_{\text{pad}} \cap \mathcal{E}_{\text{hit}}$ occurs, then for each $j \in M$, there exists a point $w_j \in \widetilde{P_j(x)}$ such that $d(x, w_j) \leq 2^{m-9}/\zeta_K$. So to get a lower bound on $d(x, W^i_t)$, we can restrict our attention to $\{w_j\}_{j \in M}$.

Claim 3.7. *If $\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \cap \mathcal{E}^\sigma_1$ occurs and*

$$d\left(w_j, \Psi_j\left(\widetilde{P_j(x)}\right)\right) \geq \varepsilon$$

for every $j \in M$, then

$$d(x, W^t_i) \geq \min\left\{ \frac{2^m}{\zeta_K}, \varepsilon - \frac{2^{m-9}}{\zeta_K} \right\}.$$

Proof. For every $j \in M$,

$$d\left(x, \Psi_j\left(\widetilde{P_j(x)}\right)\right) \geq d\left(w_j, \Psi_j\left(\widetilde{P_j(x)}\right)\right) - d(x, w_j) \geq \varepsilon - \frac{2^{m-9}}{\zeta_K}.$$

Now apply Claim 3.6. □

There are two types of points $y \in X$ which occur in the argument that follows. As a warmup, we first dispense with the easy type.

Type I: There exists $z \in B(y, 2^{m-7}/\zeta_K)$ for which $\kappa(z,t) - i \notin M$.

Fix this z and let $j' = \kappa(z,t) - i$. Assume that $\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \cap \mathcal{E}_1^\sigma$ occurs, as well as the independent event $\sigma_{j'} = 0$. Note that using Lemma 3.5 along with independence, the probability of this event is at least $q = 2^{-6} \cdot 3^{-5} \cdot (1/3)$.

Now, applying the definition of ζ_K to the sets $\widetilde{P_j(x)} = \widetilde{P_j(w_j)}$ for $j \in M$, we conclude that there is an event $\mathcal{E}_{\text{zero}}$ which occurs with probability at least p^5, and such that for every $j \in M$,

$$d\left(w_j, \Psi_j\left(\widetilde{P_j(x)}\right)\right) \geq \frac{2^{j-3}}{\zeta_K} \geq \frac{2^{m-6}}{\zeta_K}.$$

Applying Claim 3.7 with $\varepsilon = 2^{m-6}/\zeta_K$, we conclude that, in this case,

$$d(x, W_t^i) \geq \frac{5 \cdot 2^{m-9}}{\zeta_K}.$$

Since $\sigma_{j'} = 0$, we have $z \in W_t^i$, hence with probability at least $q \cdot p^5 \geq (p/16)^5$, we have

$$|d(x, W_t^i) - d(y, W_t^i)| \geq \frac{5 \cdot 2^{m-9}}{\zeta_K} - d(y,z) \geq \frac{2^{m-9}}{\zeta_K}.$$

This completes the analysis of Type I points.

3.1.3 A case analysis on the fate of y

We now analyze the complement of the set of Type I points.

Type II: For all $z \in B(y, 2^{m-7}/\zeta_K)$, $\kappa(z,t) - i \in M$.

First, we define the following key event.

$$\mathcal{E}_{\text{close}} = \left\{\exists j \in M,\ \exists z \in \widetilde{P_j(y)} \text{ such that } d(y,z) \leq \frac{2^{m-7}}{\zeta_K} \text{ and } \kappa(z,t) - i = j\right\}. \tag{7}$$

Also, let $\mathcal{E}_{\text{far}} = \neg\mathcal{E}_{\text{close}}$.

These two events concern the distance of y to the various sample sets. Since we do not make the assumption that $y \in S_m(K)$, we cannot argue that a random sample point lands near y with non-negligible probability, thus we must handle both possibilities $\mathcal{E}_{\text{close}}$ and $\mathcal{E}_{\text{far}}$. This is the main purpose of the two phases, i.e. the events $\mathcal{E}_\ell^\sigma$ for $\ell \in \{1, 2\}$. Thus the proof now breaks down into two sub-cases corresponding to the occurrences of $\mathcal{E}_{\text{close}}$ and $\mathcal{E}_{\text{far}}$, respectively.

Claim 3.8 (The close case). *Conditioned on the event $\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \cap \mathcal{E}_1^\sigma \cap \mathcal{E}_{\text{close}}$ occurring, with probability at least p^5,*

$$|d(x, W_t^i) - d(y, W_t^i)| \geq \frac{2^{m-9}}{\zeta_K}.$$

Proof. If the event $\mathcal{E}_{\text{close}} \cap \mathcal{E}_{\text{pad}}$ occurs, then there exists some $j_0 \in M$ and $z \in \widetilde{P_{j_0}(x)} = \widetilde{P_{j_0}(y)}$ such that $d(y,z) \le 2^{m-7}/\zeta_K$ and $\kappa(z,t) - 1 = j_0$. Additionally, if $\mathcal{E}_{\text{pad}} \cap \mathcal{E}_{\text{hit}}$ occurs, then for every $j \in M$ there is $w_j \in \widetilde{P_j(x)}$ with $d(w_j, x) \le 2^{m-9}/\zeta_K$. It follows that, for every $j \in M$,

$$d(w_j, z) \ge d(x,y) - d(x,w_j) - d(y,z) \ge 2^{m-1} - \frac{5 \cdot 2^{m-9}}{\zeta_K} \ge 2^{m-2} \ge 2^{j-3}.$$

Hence applying the definition of ζ_K to the sets $\widetilde{P_j(x)}$ for $j \in M$, we conclude that there is an event $\mathcal{E}_{\text{zero}}$ with probability at least p^5, independent of $\mathcal{E}_{\text{hit}}, \mathcal{E}_{\text{pad}}, \mathcal{E}_{\text{close}}$ and $\mathcal{E}_1^\sigma$, such that if the event $\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \cap \mathcal{E}_{\text{close}} \cap \mathcal{E}_1^\sigma \cap \mathcal{E}_{\text{zero}}$ occurs then $z \in \Psi_{j_0}\left(\widetilde{P_{j_0}(y)}\right)$, and for every $j \in M$,

$$d\left(w_j, \Psi_j\left(\widetilde{P_j(x)}\right)\right) \ge \frac{2^{j-3}}{\zeta_K} \ge \frac{2^{m-6}}{\zeta_K}.$$

Applying Claim 3.7, it follows that $d(x, W_t^i) \ge \frac{5 \cdot 2^{m-9}}{\zeta_K}$. Furthermore, since $z \in \Psi_{j_0}\left(\widetilde{P_{j_0}(y)}\right)$ and $\mathcal{E}_1^\sigma$ occurs, we have $\sigma_{j_0} = 1$, hence $z \in W_t^i$. It follows that

$$|d(x, W_t^i) - d(y, W_t^i)| \ge \frac{5 \cdot 2^{m-9}}{\zeta_K} - d(y,z) \ge \frac{2^{m-9}}{\zeta_K},$$

completing the proof. □

We now analyze the probability of the previous event.

Lemma 3.9. $\Pr[\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \cap \mathcal{E}_1^\sigma \,|\, \mathcal{E}_{\text{close}}] \ge 2^{-6} \cdot 3^{-5}$.

Proof. Since

$$\Pr[\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \cap \mathcal{E}_1^\sigma \,|\, \mathcal{E}_{\text{close}}] = 3^{-5} \cdot \Pr[\mathcal{E}_{\text{hit}} \cap \mathcal{E}_{\text{pad}} \,|\, \mathcal{E}_{\text{close}}] = 2^{-5} \cdot 3^{-5} \cdot \Pr[\mathcal{E}_{\text{hit}} \,|\, \mathcal{E}_{\text{pad}}, \mathcal{E}_{\text{close}}],$$

we need only argue that $\Pr[\mathcal{E}_{\text{hit}} \,|\, \mathcal{E}_{\text{pad}}, \mathcal{E}_{\text{close}}] \ge \frac{1}{2}$. But this follows by applying Lemma 2.2, part 2(a), to the sets $X = P_j(x)$, $A = B(x, 2^{m-9}/\zeta_K)$, $B = B(y, 2^{m-7}/\zeta_K)$, and concluding that

$$\Pr[\neg\mathcal{E}_{\text{hit}}^j \,|\, \mathcal{E}_{\text{pad}}, \mathcal{E}_{\text{close}}] \le e^{(1-15K)/K} \le e^{-14}.$$

Thus $\Pr[\mathcal{E}_{\text{hit}} \,|\, \mathcal{E}_{\text{pad}}, \mathcal{E}_{\text{close}}] \ge 1 - 5e^{-14} \ge \frac{1}{2}$. □

Now we proceed to analyze the case when $\mathcal{E}_{\text{far}}$ occurs. By Claim 3.4, every $w \in B(x, 2^{m-9}/\zeta_K)$ satisfies $\kappa(w,t) - i \in M$. By the pigeonhole principle, some $j^* \in M$ must occur as the value of $\kappa(w,t) - i$ in at least a $\frac{1}{5}$th of them. Together with the growth condition implied by $x \in S_m(K)$, we conclude that

$$\left|\{w \in B(x, 2^{m-9}/\zeta_K) : \kappa(w,t) - i = j^*\}\right| \ge \frac{16}{5K}|B(x, 2^{m+5}\alpha)| \ge \frac{3}{K}|B(x, 2^{m+5}\alpha)|.$$

Define the event $\mathcal{E}_{\text{hit}}^*$ to be

$$\{\exists w \in \widetilde{P_{j^*}(x)} \cap B(x, 2^{m-9}/\zeta_K) \text{ with } \kappa(w,t) - i = j^*\},$$

and observe that by Lemma 2.2, part (1), $\Pr[\mathcal{E}_{\text{hit}}^*] \ge 1 - e^{-3}$.

Claim 3.10 (The far case). *If* $\mathcal{E}_{\mathrm{pad}} \cap \mathcal{E}_{\mathrm{far}} \cap \mathcal{E}_2^\sigma \cap \mathcal{E}_{\mathrm{hit}}^*$ *occurs, then*

$$|d(x, W_t^i) - d(y, W_t^i)| \geq \frac{2^{m-9}}{\zeta_K}.$$

Proof. Assume that the event $\mathcal{E}_{\mathrm{pad}} \cap \mathcal{E}_{\mathrm{far}} \cap \mathcal{E}_2^\sigma \cap \mathcal{E}_{\mathrm{hit}}^*$ occurs, and let $w \in \widetilde{P_{j^*}(x)} \cap B(x, 2^{m-9}/\zeta_K)$ be the point guaranteed by $\mathcal{E}_{\mathrm{hit}}^*$. Since $\sigma_j = 2$, we have $w \in W_t^i$, so that $d(x, W_t^i) \leq 2^{m-9}/\zeta_K$.

On the other hand, we claim that $d(y, W_t^i) \geq 2^{m-7}/\zeta_K$. Indeed, first note that $\mathcal{E}_{\mathrm{pad}}$ implies that for all $j \in M$, $d(y, X \backslash P_j(y)) \geq 2^m$. Suppose that $z \in W_t^i$ and $d(y, z) \leq 2^{m-7}/\zeta_K$. Let $j' = \kappa(z, t) - i$. In this case, we have $j' \in M$, hence $\sigma_{j'} = 2$, and this implies that $z \in \widetilde{P_{j'}(z)} = \widetilde{P_{j'}(y)}$. But in this case, $\mathcal{E}_{\mathrm{far}}$ implies that $d(y, z) > 2^{m-7}/\zeta_K$, yielding a contradiction. It follows that

$$|d(x, W_t^i) - d(y, W_t^i)| \geq \frac{2^{m-7}}{\zeta_K} - \frac{2^{m-9}}{\zeta_K} \geq \frac{2^{m-9}}{\zeta_K}.$$

□

Lemma 3.11. $\Pr[\mathcal{E}_{\mathrm{pad}} \cap \mathcal{E}_2^\sigma \cap \mathcal{E}_{\mathrm{hit}}^* \,|\, \mathcal{E}_{\mathrm{far}}] \geq 2^{-5} \cdot 3^{-5} \cdot (1 - e^{-3})$.

Proof.

$$\Pr[\mathcal{E}_{\mathrm{pad}} \cap \mathcal{E}_2^\sigma \cap \mathcal{E}_{\mathrm{hit}}^* \,|\, \mathcal{E}_{\mathrm{far}}] = 3^{-5} \cdot 2^{-5} \cdot \Pr[\mathcal{E}_{\mathrm{hit}}^* \,|\, \mathcal{E}_{\mathrm{far}}, \mathcal{E}_{\mathrm{pad}}] \geq \Pr[\mathcal{E}_{\mathrm{hit}}^* \,|\, \mathcal{E}_{\mathrm{pad}}] \geq 3^{-5} \cdot 2^{-5} \cdot (1 - e^{-3}).$$

The penultimate inequality follows from the fact that conditioning on $\mathcal{E}_{\mathrm{far}}$ cannot decrease the probability of $\mathcal{E}_{\mathrm{hit}}^*$, as in Lemma 2.2, part 2(b). □

To finish with the analysis of the Type II points, we apply Claim 3.8 together with Lemma 3.9 and Claim 3.10 with Lemma 3.11 to conclude that

$$\begin{aligned}
\mathbb{E}|d(x, W_t^i) - d(y, W_t^i)|^2 &\geq (\Pr[\mathcal{E}_{\mathrm{good}} \cap \mathcal{E}_{\mathrm{close}} \cap \mathcal{E}_1^\sigma \cap \mathcal{E}_{\mathrm{zero}}] + \Pr[\mathcal{E}_{\mathrm{pad}} \cap \mathcal{E}_{\mathrm{far}} \cap \mathcal{E}_2^\sigma \cap \mathcal{E}_{\mathrm{hit}}^*]) \frac{2^{2m-18}}{\zeta_K^2} \\
&= (\Pr[\mathcal{E}_{\mathrm{close}}] \Pr[\mathcal{E}_{\mathrm{good}} \cap \mathcal{E}_1^\sigma \cap \mathcal{E}_{\mathrm{zero}} \,|\, \mathcal{E}_{\mathrm{close}}] + \Pr[\mathcal{E}_{\mathrm{far}}] \Pr[\mathcal{E}_{\mathrm{pad}} \cap \mathcal{E}_2^\sigma \cap \mathcal{E}_{\mathrm{hit}}^* \,|\, \mathcal{E}_{\mathrm{far}}]) \frac{2^{2m-18}}{\zeta_K^2} \\
&\geq \tfrac{1}{2} \min\{2^{-6} \cdot 3^{-5} \cdot p^5, 2^{-5} \cdot 3^{-5} \cdot (1 - e^{-3})\} \frac{2^{2m-18}}{\zeta_K^2} \\
&\geq \left(\frac{p}{64}\right)^5 \frac{2^{2m}}{\zeta_K^2}.
\end{aligned}$$

Since we have proved that (3) holds for both Type I points and Type II points, the proof is complete. □

References

[1] S. Arora, J. R. Lee, and A. Naor. Euclidean distortion and the Sparsest Cut. Manuscript, 2004.

[2] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings, and graph partitionings. In *36th Annual Symposium on the Theory of Computing*, pages 222–231, 2004.

[3] Y. Bartal. Probabilistic approximations of metric space and its algorithmic application. In *37th Annual Symposium on Foundations of Computer Science*, pages 183–193, Oct. 1996.

[4] Y. Bartal, N. Linial, M. Mendel, and A. Naor. Limitations to Fréchet's metric embedding method. *Israel J. Math.* To appear.

[5] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On metric Ramsey-type phenomena. *Ann. Math.* To appear.

[6] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.

[7] S. Chawla, A. Gupta, and H. Räcke. Embeddings of negative-type metrics and an improved approximation to generalized sparsest cut. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 102–111, Vancouver, 2005.

[8] A. Dvoretzky and C. A. Rogers. Absolute and unconditional convergence in normed linear spaces. *Proc. Nat. Acad. Sci. U. S. A.*, 36:192–197, 1950.

[9] U. Feige, M. Hajiaghayi, and J. R. Lee. Improved approximation algorithms for minimum-weight vertex separators. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 563–572, New York, NY, USA, 2005. ACM Press.

[10] A. A. Giannopoulos. A proportional Dvoretzky-Rogers factorization result. *Proc. Amer. Math. Soc.*, 124(1):233–241, 1996.

[11] M. X. Goemans. Semidefinite programming and combinatorial optimization. *Mathematical Programming*, 49:143–161, 1997.

[12] S. Khot and N. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into ℓ_1. In *Proceedings of the 46th Annual IEEE Conference on Foundations of Computer Science*, 2005. To appear.

[13] R. Krauthgamer, J. R. Lee, M. Mendel, and A. Naor. Measured descent: A new embedding method for finite metrics. *Geom. Funct. Anal.*, 2004. To appear.

[14] J. R. Lee. On distance scales, embeddings, and efficient relaxations of the cut cone. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 92–101, Vancouver, 2005.

[15] N. Linial and M. Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993.

[16] J. Matoušek and Y. Rabinovich. On dominated l_1 metrics. *Israel J. Math.*, 123:285–301, 2001.

[17] S. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the 15th Annual Symposium on Computational Geometry*, pages 300–306, New York, 1999.