Bilipschitz Snowflakes and Metrics of Negative Type

[Extended Abstract]

James R. Lee[*]
University of Washington
jrl@cs.washington.edu

Mohammad Moharrami[†]
University of Washington
mohammad@cs.washington.edu

ABSTRACT

We show that there exists a metric space (X, d) such that $(X, \sqrt{d})$ admits a bilipschitz embedding into L_2, but (X, d) does not admit an equivalent metric of negative type. In fact, we exhibit a strong quantitative bound: There are n-point subsets $Y_n \subseteq X$ such that mapping (Y_n, d) to a metric of negative type requires distortion $\widetilde{\Omega}(\log n)^{1/4}$. In a formal sense, this is the first lower bound specifically against bilipschitz embeddings into negative-type metrics, and therefore unlike other lower bounds, ours cannot be derived from a 1-dimensional Poincaré inequality.

This answers an open question about the strength of strong vs. weak triangle inequalities in a number of semi-definite programs. Our construction sheds light on the power of various notions of "dual flows" that arise in algorithms for approximating the Sparsest Cut problem. It also has other interesting implications for bilipschitz embeddings of finite metric spaces.

Categories and Subject Descriptors

F.2 [**Analysis of Algorithms and Problem Complexity**]: Geometrical problems and computations

General Terms

Algorithms, Theory

Keywords

Metric embeddings, Semi-definite programming, Sparsest Cut

1. INTRODUCTION

Beginning with the works [LLR95, AR98], it became apparent that the embeddability of finite metric spaces into various normed spaces (predominantly L_1 and L_2) was intimately tied to the efficacy of certain mathematical programs for approximating the Sparsest Cut problem in graphs. Subsequently, such tools were used to achieve new approximation results for an array of well-known problems, many of which were unapproachable via other methods.

We now recall the Sparsest Cut problem. Given a finite set V on n points, and two symmetric non-negative functions $\mathsf{cap}, \mathsf{dem} : V \times V \to \mathbb{R}_{\geq 0}$, one defines the *sparsity* of the subset $S \subseteq V$ by

$$\Phi_{\mathsf{cap},\mathsf{dem}}(S) = \frac{\mathsf{cap}(S, \bar{S})}{\mathsf{dem}(S, \bar{S})},$$

where we use the notation $f(S, \bar{S}) = \sum_{x \in S, y \notin S} f(x, y)$ for $f \in \{\mathsf{cap}, \mathsf{dem}\}$. The value of the instance $(V, \mathsf{cap}, \mathsf{dem})$ is then given by $\Phi(\mathsf{cap}, \mathsf{dem}) = \min\{\Phi_{\mathsf{cap},\mathsf{dem}}(S) : S \subseteq V\}$. We recall that the instance is said to be *uniform* if $\mathsf{dem}(u, v) = 1$ for all $u, v \in V$.

It was shown in [LLR95, AR98, GNRS99] that the integrality gap for a natural (and well-studied) linear-programming relaxation (see [LR99]) is precisely $\sup\{c_1(X, d) : (X, d)\}$, where (X, d) ranges over all metric spaces on n-points, and $c_1(X, d)$ denotes the minimal distortion required to embed (X, d) into an L_1 space (see Section 1.2 for formal definitions). Bourgain's embedding theorem [Bou85] shows that this bound is $O(\log n)$, and in [LLR95, AR98], it was shown that this is tight for the path metric on expander graphs.

The Goemans-Linial SDP. In order to achieve better approximations, one can consider the Goemans-Linial SDP:

$$\min\left\{ \frac{\sum_{u,v} \mathsf{cap}(u,v)\|x_u - x_v\|_2^2}{\sum_{u,v} \mathsf{dem}(u,v)\|x_u - x_v\|_2^2} : \{x_u\}_{u \in V} \subseteq \mathbb{R}^n \text{ and } \|\cdot\|_2^2 \text{ is a metric on } \{x_u\}_{u \in V} \right\}.$$

In other words, we optimize over sets of n vectors $W \subseteq \mathbb{R}^n$ which satisfy, for every $x, y, z \in W$, the condition

$$\|x - y\|_2^2 \leq \|x - z\|_2^2 + \|z - y\|_2^2.$$

In general, we say that a metric space (X, d) is of *negative type* if there exist a mapping $f : X \to L_2$ such that

$$\|f(x) - f(y)\|^2 = d(x, y)$$

for all $x, y \in X$.

As before (see [Mat02a, Ch. 15]), the integrality gap of this relaxation is exactly the solution to an embedding problem. The gap is precisely the supremum of $c_1(X, d)$ over all n-point metric spaces of negative type. In [ARV04], the

[*]Research supported by NSF grants CCF-0644037 and CCF-0915251, and a Sloan Research Fellowship.

[†]Research supported by NSF grant CCF-0644037.

Goemans-Linial SDP was used to achieve an $O(\sqrt{\log n})$-approximation for the *uniform* case of Sparsest Cut, and building on these techniques as well as various tools from the theory of metric embeddings, one can obtain $c_1(X,d) \le O(\sqrt{\log n}\log\log n)$ for any n-point space of negative type ([ALN08], following an earlier bound of [CGR05]). This yields the same bound for approximating the general Sparsest Cut problem.

Given the effectiveness of this approach, and generally the power of the $\|\cdot\|_2^2$ triangle inequality constraints in relaxations for other fundamental problems (see e.g. [FHL05, ACMM05, Kar09, CMM06]), it becomes a matter of fundamental importance to understand the geometry of negative-type metrics, and the effect of the negative-type constraints on mathematical programming relaxations.

Integrality gaps and the ease of snowflaking. In order to crystalize this goal, Goemans and Linial conjectured (see [Mat02a, Ch. 15], [Lin02]) that $c_1(X,d) \le O(1)$ for every metric space (X,d) of negative type. Khot and Vishnoi subsequently disproved this in [KV05]. The most ingenious part of their work involves the construction of the lower bound space (X,d), and the most intricate technical analysis goes toward showing that (X,d) is of negative type. Subsequently, [KR06] and [DKSV06] proved a stronger quantitative bound of $\Omega(\log\log n)$, where notably the latter lower bound holds in the uniform case. The first paper uses exactly the Khot-Vishnoi construction, while the second paper relies heavily on the analysis techniques of [KV05].

In [LN06], a new integrality gap construction was proposed, based on the 3-dimensional Heisenberg group $\mathbb{H}^3$. Again, the bulk of the work in [LN06] goes into proving that $\mathbb{H}^3$ admits an interesting metric of negative type. The lower bound analysis uses work of Cheeger and Kleiner [CK06b, CK06a, CK09]. Building on this analysis, it was recently proved in [CKN09] that this construction achieves an integrality gap of $(\log n)^{\delta_0}$ for some small constant $\delta_0 > 0$.

We now express a property that all these lower bounds share. For a metric space (X,d) and a number $\alpha \in (0,1]$, we use (X,d^α) to denote the metric space where X is equipped with the distance $d^\alpha(x,y) = d(x,y)^\alpha$. For values $\alpha < 1$, such constructions are commonly referred to as "snowflakes." Let us call a metric space (X,d) a *D-half-snowflake* if $c_2(X,\sqrt{d}) \le D$, i.e. $(X,\sqrt{d})$ admits a Euclidean embedding with distortion D. It is immediate from the definition that *metrics of negative type are precisely 1-half-snowflakes.*

In all the above constructions of integrality gaps, it is relatively easy to show that that space in question is an $O(1)$-half-snowflake. In the case of [KV05]-based constructions, this can be done in a page of analysis (see, e.g. [KL08]). Since the Heisenberg group $\mathbb{H}^3$ (equipped with the Carnot-Caratheodory metric) is doubling, a classical result of Assouad [Ass83] shows that it is already an $O(1)$-half-snowflake. Indeed, in all these cases, the fact that one could construct an $O(1)$-half-snowflake was taken as evidence and motivation that eventually a negative-type metric could be constructed. This leads to the following natural question (see the "Isometric vs. isomorphic L_2 squared" problem [Mat02b] posed by the first author in 2003) whose affirmative answer would make the construction of integrality gaps a significantly easier process.

Question 1. *Can every $O(1)$-half-snowflake be embedded into a metric of negative type with $O(1)$ distortion?*

The main theorem of this paper shows that the answer is negative, in a very strong sense.

Theorem 1.1. *There exists an $O(1)$-half-snowflake for which any embedding into a metric of negative type incurs distortion $\widetilde{\Omega}(\log n)^{1/4}$.*

The power of snowflakes. To give more motivation for obtaining an answer to Question 1, we observe the ubiquity of snowflakes in embeddings and analysis of the Goemans-Linial SDP and related algorithms. It is a common observation that the algorithms and analysis of [ARV04, Lee05, CGR05, ALN08] do not require the vector solution $W \subseteq \mathbb{R}^n$ to actually satisfy the full triangle inequalities, but only the weaker form: For every sequence $w_1, w_2, \ldots, w_k \in W$,

$$\|w_1 - w_k\|_2^2 \le C \sum_{i=1}^{k-1} \|w_i - w_{i+1}\|_2^2,$$

for some constant $C \ge 1$ independent of the sequence. This is merely the half-snowflake condition in disguise: It simply says that W is the image of an $O(1)$-half-snowflake embedding of some metric space. In all known algorithmic applications, it is only the $O(1)$-half-snowflake condition that is needed. Furthermore, results like [LMN05] show e.g. that planar graphs always admit $O(1)$-half-snowflake embeddings, while the question of whether they admit negative-type embeddings was posed by Rabinovich, and is still open. (This is related to another well-known conjecture in the theory of multi-commodity flows: that the shortest-path metric of a planar graph embeds into L_1 with $O(1)$ distortion [GNRS99].)

Far more than being a curiousity of the analysis, the fact that a weaker condition suffices is actually the basis for algorithmis which find sparse cuts in graphs without solving a semi-definite program. In [AHK04], the authors give an $O(\sqrt{\log n})$-approximation to the uniform Sparsest Cut problem that runs in $\widetilde{O}(n^2)$ time. In [She09], such an approximation is obtained in $\widetilde{O}(m+n^{3/2+\varepsilon})$-time for every $\varepsilon > 0$. Both of these algorithms are primal-dual, with the algorithm and analysis being guided by the structure of the Goemans-Linial SDP and its dual. A key aspect lending to their efficiency is that they do not need the full power of the dual; indeed, they operate by finding an "expander flow" [ARV04], which is a solution that corresponds precisely to a weakening of the triangle inequalities in the primal. Our lower bound shows that, in the non-uniform setting, such approaches can be significantly less powerful than finding full-fledged dual solutions. In particular, we give an $\widetilde{\Omega}(\log n)^{1/4}$ lower bound on the approximation ratio of any algorithm for non-uniform Sparsest Cut that relies on finding the non-uniform analog of expander flows. We refer to the full version for a detailed explanation.

Additional implications in metric embeddings. We mention two additional consequences of our result for the theory of bilipschitz metric embeddings, which have played a fundamental role in approximation algorithms.

Scale gluing for L_1. The first application is to "scale-gluing" results for embeddings into L_1. Specifically, suppose that (X,d) is an n-point metric space, and furthermore that for every value $k \in \mathbb{Z}$, there exists a 1-Lipschitz map $\varphi_k : X \to L_1$ such that for $x, y \in X$ satisfying $d(x,y) \geq 2^k$, we have

$$\|\varphi_k(x) - \varphi_k(y)\|_1 \geq \frac{2^k}{\alpha}. \quad (1)$$

On the one hand, we have the following.

THEOREM 1.2. *[Lee05] There exists an embedding of (X,d) into L_1 with distortion $O(\sqrt{\alpha \log n})$.*

Such scale-gluing results have played a central role in the best-known approximation algorithms for Sparsest Cut [KLMN05, Lee05, CGR05, ALN08] and Graph Bandwidth [Rao99, KLMN05, Lee09], and have found applications in approximate multi-commodity max-flow/min-cut theorems in graphs [Rao99, KLMN05]. The work of Cheeger and Kleiner [CK06a] shows that even if $\alpha = O(1)$, the distortion can go to infinity, and [CKN09] gives a definite bound of $\Omega(\log n)^{\delta_0}$ for some small $\delta_0 > 0$. Our construction and analysis yield the following lower bound.

THEOREM 1.3. *There exist n-point metric spaces (X,d) which satisfy (1) with $\alpha = 2$, but such that $c_1(X,d) = \widetilde{\Omega}\,(\log n)^{1/4}$. In fact, we even have $c_{\mathsf{NEG}}(X,d) = \widetilde{\Omega}\,(\log n)^{1/4}$.*

The k-sum embedding conjecture. In [LS09], it is conjectured that if a family of finite graphs $\mathcal{F}$ is such that every shortest-path metric supported on a member of $\mathcal{F}$ embeds into L_1 with distortion $O(1)$, then the family $\oplus_k \mathcal{F}$ has the same property, for every $k \in \mathbb{N}$, where the $\oplus_k(\cdot)$ notation denotes the closure of $\mathcal{F}$ under the operation of taking k-sums along cliques (see [LS09] for a formal description). This authors refer to this as the "k-sum embedding conjecture." The conjecture is open even for $k = 2$. One of the main results of [LS09] is that the k-sum embedding conjecture, combined with the well-known planar embedding conjecture, implies the GNRS max-flow/min-cut conjecture in excluded-minor families [GNRS99].

Our results show that there exists an unweighted graph G whose shortest-path metric embeds into L_1 with distortion 2, but that by taking repeated 2-sums of G with itself, one obtains a graph whose L_1 distortion becomes arbitrarily large. This does not disprove the k-sum conjecture, because we have only considered a single shortest-path metric on G, but it does show that the proof must use something about the entire set of embeddings for metrics on G, as opposed to merely an embedding of the given metric on G.

1.1 Efficiency, embeddings, and iterated graphs

We now describe our main construction, and the steps that go into a exhibiting a separation between half-snowflakes and metrics of negative type. Various aspects are simplified here for the sake of exposition, and some finer points of the construction are left unmentioned. The analysis is based on a "differentiation"-type argument. At a very broad level, we first argue that any low-distortion embedding must be well-controlled on a small piece of our lower bound space, and then show that any well-controlled embedding is quite rigid in structure, allowing us to prove a lower bound.

Generalizations of classical differentiation theory have played a prominent role in proving the non-existence of bi-Lipschitz embeddings between various spaces, when the target space Z is sufficiently nice (e.g. if Z is a Banach space with the

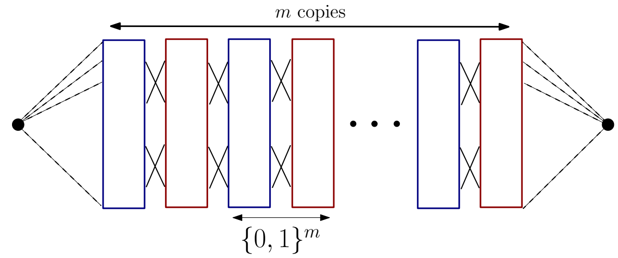

Figure 1: The "string of cubes" graph, which we recursively compose with itself.

Radon-Nikodym property); see, for instance [Pan89, Che99, LN06, BL00, CK06c]. But this approach does not apply to targets like L_1 which don't have the Radon-Nikodym property; in particular, even Lipschitz mappings $f : \mathbb{R} \to L_1$ are not guaranteed to be differentiable in the classical sense.

More recently, however, Cheeger and Kleiner [CK06a, CK06b] have successfully applied weaker notions of differentiability to the study of L_1 embeddings of the Heisenberg group. Subsequent papers [LR07, CK09, CKN09] continue this theme. Ours is the first work to apply these techniques to lower bounds against negative-type metrics.

The construction. Let G be an unweighted graph with two distinguished vertices $s, t \in V(G)$. As in [LR07], we use $G^{\oslash k}$ to denote the following iterated graph: $G^{\oslash 0}$ is a single edge, and $G^{\oslash k+1}$ arises by replacing every edge of $G^{\oslash k}$ with a copy of G, with s and t taking the place of the endpoints of the edge. See Section 2 for a formal definition.

For a parameter $m \in \mathbb{N}$, consider now the graph H_m constucted as follows. Let Q_m be the m-dimensional hypercube graph, and write $V(Q_m) = B_m \cup R_m$, where B_m and R_m denote the nodes of even and odd parity, respectively. Then Q_m is bipartite with respect to the partition (B_m, R_m). H_m is the graph which consists of $2m$ layers of the form

$$B_m^{(1)} R_m^{(1)} B_m^{(2)} R_m^{(2)} B_m^{(3)} R_m^{(3)} \cdots B_m^{(m)} R_m^{(m)}, \quad (2)$$

where $B_m^{(i)}$ and $R_m^{(i)}$ denote disjoint copies of B_m and R_m for $i = 1, 2, \ldots, m$, and hypercube edges are presently between every pair of adjacent layers.

We also add to H_m two distinguished nodes: s connected to all the nodes of $B_m^{(1)}$ by paths of length m and a t connected the nodes of $R_m^{(m)}$ by paths of length t. We call H_m the "string of cubes" graph. See Figure 1. Our final construction is of the form $G_{k,m} = H_m^{\oslash k}$ for appropriate values of $k, m \in \mathbb{N}$. We use $d_{k,m}$ to denote the shortest-path metric on $G_{k,m}$. Our goal is now to show that $G_{k,m}$ is an $O(1)$-half-snowflake for all $k, m \in \mathbb{N}$, while $\lim_{m\to\infty} \lim_{k\to\infty} c_{\mathsf{NEG}}(G_{k,m}) = \infty$.

Efficiency. A central role will be played by the *efficiency* of various "paths" in metric spaces, following [EFW06, LR07]. Consider a finite sequence of points equipped with a non-negative symmetric function dist (which may *not* satisfy the triangle inequality), $\mathcal{S} = \{x_1, x_2, \ldots, x_k\}$. We say that $\mathcal{S}$ is *ε-efficient (with respect to* dist*)* if

$$\sum_{i=1}^{k-1} \mathsf{dist}(x_i, x_{i+1}) \leq (1+\varepsilon)\,\mathsf{dist}(x_1, x_k).$$

Note that if dist is a metric, then the left-hand side is always at least $\mathsf{dist}(x_1, x_k)$, by the triangle inequality.

The first key aspect of our approach is that we reduce the embeddability of $G^{\oslash k}$ to the study of specific types of embeddings for the base graph G. For a graph G and two nodes $s, t \in V(G)$, let $\mathcal{P}_{s,t}(G)$ denote the set of all s-t shortest-paths in G. In Section 2, we prove a quantitative variant of the following theorem, based on the "coarse differentiation" methodology of [EFW06].

THEOREM 1.4. *Let (Y, d_Y) be any metric space, and suppose that $c_Y(G^{\oslash k}) \le D$ for all $k = 1, 2, \ldots$. Then for every $\varepsilon > 0$, there exists an embedding $f : G \to Y$ with distortion at most D, and such that for every sequence $\{x_1, x_2, \ldots, x_r\} \in \mathcal{P}_{s,t}(G)$, the sequence $\{f(x_1), f(x_2), \ldots, f(x_r)\}$ is ε-efficient in (Y, d_Y).*

A more general result was proved in [LR07]. The novel aspect of our approach in Section 2, is that in the special case of iterated graphs, we are able to gain an exponential improvement in a quantitative version of this theorem (which is necessary to obtain our strong lower bounds). In particular, the dependence of k on D is logarithmic (as opposed to linear as in [LR07]).

Snowflake embedding: On the other hand, in the setting of half-snowflakes, we have a partial converse to the preceding theorem. If there is an embedding

$$f : (G, \sqrt{d_G}) \to L_2$$

with distortion D, and such that for every sequence

$$\{x_1, x_2 \ldots, x_r\} \in \mathcal{P}_{s,t}(G),$$

the sequence $\{f(x_1), f(x_2), \ldots, f(x_r)\}$ is ε-efficient with respect to the distance $\|f(x_i) - f(x_j)\|_2^2$, then for every $k = O(1/\varepsilon)$, the graph $G^{\oslash k}$ is an $O(D)$-half-snowflake. (We are actually only able to prove this for a modification of the graph $G^{\oslash k}$.) Due to lack of space, the proof of this fact is deferred to the full version.

Because of these two results, we are able to focus on a separation between embeddings of our base graph H_m into negative-type metrics and half-snowflakes, respectively, with the additional property that the embeddings are ε-efficient on s-t shortest-paths. Analyzing efficient (and approximately efficient embeddings) is the technical core of our approach, which we now address.

A lower bound for efficient embeddings into NEG: Poincaré boosting. Recall that H_m consists of m hypercubes $C_1, C_2, \ldots, C_m$ strung together as in (2). First, we recall the classical Poincaré inequality of Enflo [Enf69] for the discrete m-cube. For any $f : Q_m \to \mathbb{R}$, we have

$$\mathbb{E}_{x \in Q_m} |f(x) - f(\bar{x})|^2 \le \sum_{i=1}^{m} \mathbb{E}_{x \in Q_m} |f(x) - f(x \oplus e_i)|^2, \quad (3)$$

where we use $\bar{x}$ to denote x with all coordinates flipped, and $x \oplus e_i$ to denote x with the ith coordinate flipped. By integrating, we easily conclude that for any $f : Q_m \to L_2$,

$$\mathbb{E}_{x \in Q_m} \|f(x) - f(\bar{x})\|^2 \le \sum_{i=1}^{m} \mathbb{E}_{x \in Q_m} \|f(x) - f(x \oplus e_i)\|^2. \quad (4)$$

Obviously, this inequality does not yield any lower bound on the distortion for embedding Q_m into NEG (since it embeds into L_1, and hence NEG isometrically).

But suppose we are given an embedding $g : H_m \to L_2$ which is an isometric negative-type embedding, in the sense that $\|g(x) - g(y)\|^2 = d_{H_m}(x, y)$ for all $x, y \in H_m$. Now, for each i, let $g_i = g|_{C_i} : Q_m \to L_2$ be the restriction of g to the ith copy of Q_m in H_m, where we think of all the maps $\{g_i\}_{i=1}^m$ as having the same domain. If we simply applied (4) to each $f = g_i$ and summed the resulting inequalities, we would again achieve no non-trivial lower bound.

Instead, we apply (4) to the mapping $f = g_1 + g_2 + \cdots + g_m$. By the strictness of the property that g is an isometry (when the range is considered with the squared norm $\|\cdot\|^2$), all the vectors $\{g_i(x) - g_i(\bar{x})\}_{i=1}^m$ are colinear, and one concludes that,

$$\|f(x) - f(\bar{x})\|^2 = \left(\sum_{j=1}^{m} \|g_j(x) - g_j(\bar{x})\| \right)^2. \quad (5)$$

On the other hand, if we (by abuse of notation) consider a shortest-path in H_m of the form

$$x \,-\, x \oplus e_i \,-\, x \,-\, x \oplus e_i \,-\, x \,-\, x \oplus e_i \,-\, \cdots$$

(where the elements of the path lie in the respective sets $B_m^{(1)}, R_m^{(1)}, B_m^{(2)}, R_m^{(2)}, \ldots$), then by the fact that g is an isometry, for any pairs of adjacent nodes x, x' and y, y' in such a path, we have $g(x) - g(x')$ and $g(y) - g(y')$ being orthogonal. This implies that for every $x \in Q_m$ and $i \in [m]$, we have for every $j, k \in [m]$, the property that $g_j(x) - g_j(x \oplus e_i)$ and $g_k(x) - g_k(x \oplus e_i)$ are orthogonal (actually, this only holds for j and k of the same parity, but ignore this small issue).

We conclude that if $f = g_1 + g_2 + \cdots g_m$, then

$$\|f(x) - f(x \oplus e_i)\|^2 = \sum_{j=1}^{m} \|g_j(x) - g_j(x \oplus e_i)\|^2. \quad (6)$$

From (5), (6), and (4), we get a "boosted" Poincaré inequality of the form

$$\mathbb{E}_{x \in Q_m} \left(\sum_{j=1}^{m} \|g_j(x) - g_j(\bar{x})\| \right)^2 \quad (7)$$
$$\le \sum_{j=1}^{m} \sum_{i=1}^{m} \mathbb{E}_{x \in Q_m} \|g_j(x) - g_j(x \oplus e_i)\|^2.$$

Notice that to obtain this inequality, we needed to use the fully high-dimensional version (4) instead of (3), because we used the high-dimensional relationship between the various maps $\{g_i\}_{i=1}^m$.

Now, if we were told in advance that (7) holds, and also that each g_j has distortion at most D (again, when the range is considered with the squared distance $\|\cdot\|^2$), it would immediately yield a lower bound of $D \ge m$. (Assuming each g_j is 1-Lipschitz, the right-hand side is at most m^2, while the left-hand side is at least m^3/D.)

Of course, we started with the assumption that g was isometric, so this simply proves that H_m does not admit an isometric negative-type embedding. But now the main point is that every aspect of the preceding argument is *robust.* In Section 3, we prove a stable version of (5) using a distortion bound for g, and a stable version of (6) using the assumption that g is ε-efficient on a large fraction of s-t shortest-paths in H_m.

Combining all this together in Section 3 shows that H_m does not admit a low-distortion negative-type metric which is ε-efficient on most s-t paths (for ε small enough). Combined with a differentiation theorem like Theorem 1.4, this shows that the iterated graph $G_{k,m}$ does not admit a low-distortion negative-type metric for k, m large enough.

An upper bound for efficient embeddings into half-snowflakes. The preceding discussion yields only part of the separation between half-snowflakes and negative-type metrics. For the other side, using the "snowflake embedding" method mentioned earlier for the iterated graph $G_{k,m}$, it suffices to construct an embedding $f : (H_m, \sqrt{d_{H_m}}) \to L_2$ which has small distortion, and such that every s-t shortest-path in H_m is mapped 0-efficiently, when the range is considered with the squared distance $\|\cdot\|^2$. To illustrate how this is done, we will argue for the path metric P on the points $\{1, 2, \ldots, n\}$. In the actual construction, the following argument is carried out for all the shortest s-t paths simulataneously.

Suppose that $f : P \to L_2$ satisfies, for all $x, y \in P$,

$$\frac{|x-y|}{D} \le \|f(x) - f(y)\|^2 \le |x-y|. \tag{8}$$

Let $v_0 = \frac{f(n)-f(1)}{\|f(n)-f(1)\|}$, and put $\alpha_i = \langle v_0, f(i)\rangle$. We will also need to make the assumption that

$$\alpha_1 \le \alpha_2 \le \cdots \le \alpha_n, \tag{9}$$

which will be satisfied in our constructions.

Now, write every point $i \in P$ in the form

$$f(i) = \alpha_i v_0 + v_i,$$

where $\langle v_i, v_0\rangle = 0$. Consider the mapping $g(i) = \alpha_i v_0 + \delta v_i$, for some $\delta \in [0, 1]$.

In this case, we have

$$\begin{aligned}\sum_{i=1}^{n-1} \|g(i+1) - g(i)\|^2 &= \sum_{i=1}^{n-1} (\alpha_{i+1} - \alpha_i)^2 + \delta^2 \|v_{i+1} - v_i\|^2 \\ &\le \sum_{i=1}^{n-1} (\alpha_{i+1} - \alpha_i)^2 + \delta^2 n,\end{aligned}$$

where in the last inequality we have used (8).

Note that $\|f(n) - f(1)\| = \sum_{i+1} |\alpha_{i+1} - \alpha_i|$ by (9), and we have $|\alpha_{i+1} - \alpha_i| \le 1$ by (8), hence

$$\sum_{i=1}^{n-1} (\alpha_{i+1} - \alpha_i)^2 \le \|f(n) - f(1)\|.$$

On the other hand, $\|g(n) - g(1)\|^2 = \|f(n) - f(1)\|^2 \ge n/D$. It follows that for some value $\delta \gtrsim 1/\sqrt{D}$, we will have

$$\|g(n) - g(1)\|^2 = \sum_{i=1}^{n-1} \|g(i+1) - g(i)\|^2,$$

i.e. the image will of P will be 0-efficient. This gives us a general way to obtain 0-efficient embeddings for half-snowflakes, which would not work for negative-type embeddings (because in the process of decreasing δ, triangle inequalities that may have been satisfied in the image could become violated).

1.2 Preliminaries

For a graph G, we will use $V(G), E(G)$ to denote the sets of vertices and edges of G, respectively. Sometimes we will equip G with a non-negative length function $\mathsf{len} : E(G) \to \mathbb{R}_+$, and we let d_{len} denote the shortest-path pseudo-metric on G. We refer to the pair (G, len) as a *metric graph,* and often len will be implicit, in which case we use d_G to denote the path metric.

Given two expressions E and E' (possibly depending on a number of parameters), we write $E = O(E')$ to mean that $E \le CE'$ for some constant $C > 0$ which is independent of the parameters. Similarly, $E = \Omega(E')$ implies that $E \ge CE'$ for some $C > 0$. We also write $E \lesssim E'$ as a synonym for $E = O(E')$. Finally, we write $E \asymp E'$ to denote the conjunction of $E \lesssim E'$ and $E \gtrsim E'$.

Embeddings and distortion. If $(X, d_X), (Y, d_Y)$ are metric spaces, and $f : X \to Y$, then we write

$$\|f\|_{\mathrm{Lip}} = \sup_{x \ne y \in X} \frac{d_Y(f(x), f(y))}{d_X(x, y)}.$$

If f is injective, then the *distortion of* f is defined by $\mathsf{dist}(f) = \|f\|_{\mathrm{Lip}} \cdot \|f^{-1}\|_{\mathrm{Lip}}$. A map with distortion D will sometimes be referred to as D*-bi-lipschitz.* If $d_Y(f(x), f(y)) \le d_X(x, y)$ for every $x, y \in X$, we say that f is *non-expansive.* If $d_Y(f(x), f(y)) \ge d_X(x, y)$ for every $x, y \in X$, we say that f is *non-contracting.* For a metric space X, we use $c_p(X)$ to denote the least distortion required to embed X into some L_p space.

A metric space (X, d) is said to be of *negative type* if $c_2(X, \sqrt{d}) = 1$. We use $c_{\mathsf{NEG}}(X, d)$ to denote the least distortion required to embed X into some metric space of negative type. We will abuse notation in the following way. We write $f : (X, d) \to \mathsf{NEG}$ to denote the fact that f takes values in L_2, and that its image is a metric when equipped with the distance induced from the square of the L_2 norm. In this case, the image is always thought to be equipped with the metric $\|f(x) - f(y)\|_2^2$, for $x, y \in X$ (e.g. for notions like the distortion of f).

2. ITERATED GRAPHS AND COARSE DIFFERENTIATION

In the present section, we define formally a key part of our graph construction process (marked $\oslash$ products), and give a differentiation theorem for iterated product graphs.

2.1 Marked $\oslash$-products

An s-t graph G is a graph which has two distinguished vertices $s, t \in V(G)$. For an s-t graph, we use $s(G)$ and $t(G)$ to denote the vertices labeled s and t, respectively. We define the length of an s-t graph G as $\mathsf{len}(G) = d_{\mathsf{len}}(s, t)$. Throughout the paper, we will only be concerned with *symmetric* s-t graphs, i.e. graphs for which there is an automorphism which maps s to t. We assume that all s-t graphs are symmetric in the following definitions. A *marked graph* $G = (V, E)$ is one which carries an additional subset $E_M(G) \subseteq E$ of *marked edges.* Every graph is assumed to be equipped with the trivial marking $E_M(G) = E(G)$ unless a marking is otherwise specified.

Definition 2.1 (Composition of s-t graphs). *Given two marked s-t graphs H and G, define $H \oslash G$ to be the s-t graph obtained by replacing each* marked *edge $(u, v) \in E_M(H)$ by a copy of G. Formally,*

- $V(H \oslash G) = V(H) \cup (E_M(H) \times (V(G) \setminus \{s(G), t(G)\}))$.

- *For every edges $(u,v) \in E(H) \setminus E_M(H)$, there is a corresponding edge in $H \oslash G$.*
- *For every edge $e = (u,v) \in E_M(H)$, there are $|E(G)|$ edges,*

$$\Big\{\Big((e,v_1),(e,v_2)\Big) \mid (v_1,v_2) \in E(G), v_1,v_2 \notin \{s(G),t(G)\}\Big\}$$
$$\cup \Big\{\Big(u,(e,w)\Big) \mid (s(G),w) \in E(G)\Big\}$$
$$\cup \Big\{\Big((e,w),v\Big) \mid (w,t(G)) \in E(G)\Big\}$$

- *The marked edges of $H \oslash G$ are precisely those introduced in the previous step which correspond to marked edges in G.*
- *$s(H \oslash G) = s(H)$ and $t(H \oslash G) = t(H)$.*

If H and G are equipped with length functions $\mathsf{len}_H, \mathsf{len}_G$, respectively, we define $\mathsf{len} = \mathsf{len}_{H\oslash G}$ as follows. Using the preceding notation, for every edge $e = (u,v) \in E_M(H)$,

$$\begin{aligned}
\mathsf{len}\left((e,v_1),(e,v_2)\right) &= \frac{\mathsf{len}_H(e)}{d_{\mathsf{len}_G}(s(G),t(G))}\mathsf{len}_G(v_1,v_2)\\
\mathsf{len}\left(u,(e,w)\right) &= \frac{\mathsf{len}_H(e)}{d_{\mathsf{len}_G}(s(G),t(G))}\mathsf{len}_G(s(G),w)\\
\mathsf{len}\left((e,w),v\right) &= \frac{\mathsf{len}_H(e)}{d_{\mathsf{len}_G}(s(G),t(G))}\mathsf{len}_G(w,t(G)).
\end{aligned}$$

This choice implies that $H \oslash G$ contains an isometric copy of $(V(H), d_{\mathsf{len}_H})$.

Definition 2.2 (Recursive composition). *Given a marked s-t graph G and a number $k \in \mathbb{N}$, we define $G^{\oslash k}$ inductively by letting $G^{\oslash 0}$ be a single edge of unit length, and setting $G^{\oslash k} = G^{\oslash k-1} \oslash G$.*

The following result is straightforward.

Lemma 2.3 (Associativity of $\oslash$). *For any three graphs A, B, C, we have $(A \oslash B) \oslash C = A \oslash (B \oslash C)$, both graph-theoretically and as metric spaces.*

Definition 2.4. *For two graphs G, H, a subset of vertices $X \subseteq V(H)$ is said to be a copy of G if there exists a bijection $f : V(G) \to X$ with distortion 1.*

Now we make the following two simple observations about copies of H and G in $H \oslash G$.

Observation 2.5. *The graph $H \oslash G$ contains $|E_M(H)|$ distinguished copies of the graph G, one copy corresponding to each edge in H.*

Observation 2.6. *The subset of vertices $V(H) \subseteq V(H \oslash G)$ form an isometric copy of H.*

2.2 Coarse differentiation

Let G be an s-t graph, (X,d) a metric space, and consider a mapping $f : V(G) \to X$. Recalling that $\mathcal{P}_{s,t}(G)$ is the set of s-t shortest-paths in G, let μ be a measure on $\mathcal{P}_{s,t} = \mathcal{P}_{s,t}(G)$. We say that f is *ε-efficient with respect to μ* if it satisfies

$$\mathbb{E}_{\gamma\sim\mu} \sum_{uv\in\gamma} d(f(u),f(v)) \le (1+\varepsilon)\, d(f(s),f(t)).$$

For a marked s-t graph G, we define its *marked length* by

$$\mathsf{len}_M(G) = \min_{\gamma\in\mathcal{P}_{s,t}} \sum_{uv\in\gamma:(u,v)\in E_M(G)} \mathsf{len}_G(u,v).$$

Theorem 2.7. *Let G be a marked s-t graph. Then for any $D \ge 1$ and $\varepsilon \ge 2D\left(1 - \frac{\mathsf{len}_M(G)}{\mathsf{len}(G)}\right)$, there exists a $k = O(\frac{1}{\varepsilon}\log D)$ such that the following holds. For every metric space (X,d) and mapping $f : V(G^{\oslash k}) \to X$ with distortion D, there exists a copy of G in $G^{\oslash k}$ such that $f|_G$ is ε-efficient with respect to μ.*

Proof. Assume, without loss of generality, that f is 1-Lipschitz. We claim that if f is not ε-efficient with respect to μ on any copy of G in $G^{\oslash k}$, then

$$\begin{aligned}
&\mathbb{E}_{\gamma\sim\mu^{\oslash k}} \sum_{uv\in\gamma} d(f(u),f(v)) \qquad (10)\\
&\qquad \ge \left(1+\frac{\varepsilon}{2}\right)^k d\left(f(s(G^{\oslash k})), f(t(G^{\oslash k}))\right),
\end{aligned}$$

where $\mu^{\oslash k}$ is the natural iterated measure on $s(G^{\oslash k})$-$t(G^{\oslash k})$ shortest-paths in $G^{\oslash k}$. We prove this by induction on k, where the case $k=0$ is trivial.

Now, write $G^{\oslash k+1} = G \oslash G^{\oslash k}$. We have one copy of $G^{\oslash k}$ in $G^{\oslash k+1}$ for every marked edge $e = (u,v) \in E_M(G)$. Denoting this copy by H_e, by the induction hypothesis, we have

$$\begin{aligned}
&\mathbb{E}_{\gamma\sim(\mu^{\oslash k})_e} \sum_{uv\in\gamma} d(f(u),f(v)) \qquad (11)\\
&\qquad \ge \left(1+\frac{\varepsilon}{2}\right)^k d\left(f(s(H_e)), f(t(H_e))\right),
\end{aligned}$$

where we use $(\mu^{\oslash k})_e$ to denote the path measure on H_e. Denote now the outer copy of G by G_0, and observe that if it is not mapped ε-efficiently by f, then

$$\begin{aligned}
\mathbb{E}_{\gamma\sim\mu} \sum_{uv\in\gamma} d(f(u),f(v)) &\ge (1+\varepsilon)\, d(f(s(G_0)), f(t(G_0)))\\
&= (1+\varepsilon)\, d\left(f(s(G^{\oslash k+1})), f(t(G^{\oslash k+1}))\right),
\end{aligned}$$

where the distribution μ here is over $s(G_0)$-$t(G_0)$ paths in $G_0 \cong G$.

Consider now a term of the form $\sum_{uv\in\gamma} d(f(u),f(v))$ on the left-hand side. Since f is 1-Lipschitz, we have

$$\begin{aligned}
&\sum_{uv\in\gamma:(u,v)\in E_M(G_0)} d(f(u),f(v))\\
&\ge \sum_{uv\in\gamma} d(f(u),f(v)) - \left(1 - \frac{\mathsf{len}_M(G_0)}{\mathsf{len}(G_0)}\right) d_G(s(G_0),t(G_0))\\
&\ge \sum_{uv\in\gamma} d(f(u),f(v)) - \frac{\varepsilon}{2D}\, d(f(s(G_0)), f(t(G_0)))\\
&\ge \left(1+\frac{\varepsilon}{2}\right) d(f(s(G_0)), f(t(G_0))).
\end{aligned}$$

But now since each marked edge is replaced by a copy of $G^{\oslash k}$ in G, every term in the sum $\sum_{uv\in\gamma:(u,v)\in E_M} d(f(u),f(v))$ corresponds to the right-hand side of an instance of (11), yielding

$$\begin{aligned}
&\mathbb{E}_{\gamma\sim\mu^{\oslash k+1}} \sum_{uv\in\gamma} d(f(u),f(v)) \qquad (12)\\
&\qquad \ge \left(1+\frac{\varepsilon}{2}\right)^{k+1} d\left(f(s(G^{\oslash k+1})), f(t(G^{\oslash k+1}))\right).
\end{aligned}$$

This preceding line completes our proof of (10) by induction. Now, combining the triangle inequality and the fact that f is 1-Lipschitz, the left-hand side of (10) is at most $\mathsf{len}(G^{\oslash k})$, while the right-hand side is at least

$$\frac{\mathsf{len}(G^{\oslash k})}{D}\left(1+\frac{\varepsilon}{2}\right)^k,$$

yielding a contradiction for $k \asymp \frac{\log D}{\varepsilon}$. $\square$

3. LOWER BOUND FOR NEG

We now formally present our lower bound construction and prove a lower bound on the distortion required to embed these graphs into NEG.

3.1 Graph construction

We will refer to the graphs H_m described in Section 1.1. We use $[Q_m]_i$ to denote the ith copy of Q_m in H_m, and for a vertex $x \in V(Q_m)$, we use $[x]_i$ to denote the copy of x in $[Q_m]_i$. For a directed edge $\vec{e} = (u, v)$, we define $f(\vec{e}) = f(u) - f(v)$.

For $m, h \in \mathbb{N}$, we define the graph $I_{m,h}$ as follows. We begin with a copy of H_m where all the edges are marked. Then, we relabel the vertices s and t in H_m as s' and t'. Next, we add two distinguished vertices s and t and and connect s to s' and t to t' by a path of length $1000\lceil\sqrt{m}\rceil$. All the edges in this new path are unmarked. Finally, we replace each marked edge with a path of length $1000\lceil\sqrt{m}\rceil h$, all of whose edges are marked. Our final construction is of the form $I_{m,h}^{\oslash k}$ for appropriate choices of $m, h, k \in \mathbb{N}$. We equip these graphs with the unweighted shortest-path metric, which we denote $d_{m,h,k}$.

3.2 Distortion lower bound

The main result of this section is the following theorem.

THEOREM 3.1. *For $m \geq 1$ and $k = \lceil m \log^2 m\rceil$, any embedding of $I_{m,m^2}^{\oslash k}$ into NEG requires distortion $\tilde{\Omega}(\log^{1/4} N)$ where $N = |V(H_m^{\oslash k})| \asymp 2^{O(mk)}$.*

We will use Theorem 2.7 to reduce our task to proving lower bounds on efficient embeddings. Let μ_m be the uniform measure over s-t shortest paths in H_m of the form

$$(s, \dots, [x]_1, [x \oplus e_k]_1, [x]_2, [x \oplus e_k]_2, \dots, [x]_m, [x \oplus e_k]_m, \dots, t),$$

where $k \in \{1, 2, \dots, m\}$ and $x \in Q_m$ are chosen uniformly at random.

LEMMA 3.2. *For any mapping $f : H_m \to$ NEG, if f is $O(\frac{1}{m \log m})$-efficient with respect to μ_m, then $\mathsf{dist}(f) \gtrsim m^{1/2}$.*

Using the preceding lemma, the Theorem 3.1 follows quickly.

PROOF OF THEOREM 3.1. Suppose that $f : I_{m,m^2}^{\oslash k} \to$ NEG has distortion at most $\sqrt{m}$. By Theorem 2.7, there must exist an isometric copy of $(I_{m,m^2}, d_{m,m^2,1})$, in $I_{m,m^2}^{\oslash k}$ such that $f|_{I_{m,m^2}}$ is $\frac{1}{m\log m}$-efficient with respect to μ_m. Now, I_{m,m^2} possesses an isometric copy of (H_m, d_{H_m}) as a subgraph.

We restrict f further to H_m. The map $f|_{H_m} : H_m \to$ NEG is $\frac{1}{m \log m} + O(1/m^2)\mathsf{dist}(f)$-efficient, and is thus $O(\frac{1}{m\log m})$-efficient. By Lemma 3.2, any such embedding of H_m has distortion $\Omega(\sqrt{m})$. We have $\log N \lesssim m^2 \log^2 m$, hence $\mathsf{dist}(f) \gtrsim m^{1/2} \gtrsim \frac{(\log N)^{1/4}}{\log\log N}$. $\square$

Before we prove Lemma 3.2, we start with a couple of general lemmas. Let $\vec{E}(Q_m)$ be the set of all ordered pairs (u, v) where $\{u, v\}$ is an edge of Q_m (in other words, replace every undirected edge by two directed edges).

LEMMA 3.3. *Let $\mathcal{F}$ be a set of functions $f : Q_m \to L_2$ such that*

i) $\mathbb{E}_{f,g\in\mathcal{F},\vec{e}\in\vec{E}(Q_m)}[f(\vec{e})\cdot g(\vec{e})] \leq c_1$,

ii) $\mathbb{E}_{f,g\in\mathcal{F},x\in V(Q_m)}[(f(x) - f(\bar{x}))\cdot(g(x) - g(\bar{x}))] \geq c_2$.

Then,

$$m \geq \frac{c_2}{c_1}.$$

PROOF. Let $F(x) = \mathbb{E}_{f\in\mathcal{F}} f(x)$. Therefore,

$$\begin{aligned} c_2 &\leq \mathbb{E}_{f,g\in\mathcal{F},x\in V(Q_m)}[(f(x) - f(\bar{x}))\cdot(g(x) - g(\bar{x}))] \\ &= \mathbb{E}_{x\in V(Q_m)}[\mathbb{E}_{f\in\mathcal{F}}(f(x) - f(\bar{x}))\cdot \mathbb{E}_{g\in\mathcal{F}}(g(x) - g(\bar{x}))] \\ &= \mathbb{E}_{x\in V(Q_m)}[(F(x) - F(\bar{x}))\cdot(F(x) - F(\bar{x}))] \\ &= \mathbb{E}_{x\in V(Q_m)}[\|F(x) - F(\bar{x})\|_2^2]. \end{aligned}$$

Therefore (4) implies,

$$\begin{aligned} c_2 &\leq m\cdot \mathbb{E}_{\vec{e}\in\vec{E}(Q_m)}[\|F(\vec{e})\|_2^2] \\ &= m\cdot \mathbb{E}_{\vec{e}\in\vec{E}(Q_m)}[(F(\vec{e}))\cdot(F(\vec{e}))] \\ &= \mathbb{E}_{\vec{e}\in\vec{E}(Q_m)}[\mathbb{E}_{f\in\mathcal{F}} f(\vec{e})\cdot \mathbb{E}_{g\in\mathcal{F}} g(\vec{e})] \\ &= \mathbb{E}_{f,g\in\mathcal{F},\vec{e}\in\vec{E}(Q_m)}[f(\vec{e})\cdot g(\vec{e})] \\ &= m\cdot c_1 \end{aligned}$$

$\square$

The main idea in the proof of Lemma 3.2 is to apply Lemma 3.3 to some family $\mathcal{F} \subseteq \{f|_{[Q_m]_i} : 1 \leq i \leq m\}$. In the rest of the section, we are mainly concerned with bounding the values c_1 and c_2 based on efficiency and distortion.

LEMMA 3.4. *For all $i, j \in [m]$, and for any non-expansive map $f : V(H_m) \to$ NEG with distortion at most D, the following inequality holds,*

$$(f([x]_i) - f([\bar{x}]_i))(f([x]_j) - f([\bar{x}]_j)) \geq \frac{m}{D} - 4(|j-i|). \quad (13)$$

PROOF. We prove this inequality using the following bounds.

i) Non-expanding property: $\|f([x]_j) - f([x]_i)\|_2^2 \leq 2|i-j|$ and $\|f([\bar{x}]_i) - f([\bar{x}]_j)\|_2^2 \leq 2|i-j|$.

ii) Distortion bound: $\|f([x]_i) - f([\bar{x}]_i)\|_2^2 \geq \frac{m}{D}$ and $\|f([x]_j) - f([\bar{x}]_j)\|_2^2 \geq \frac{m}{D}$.

To prove (13),

$$\begin{aligned} &\frac{m}{D} - 4|i-j| \\ &\leq \frac{m}{D} - \|(f([x]_i) - f([x]_j))\|_2^2 - \|(f([\bar{x}]_j) - f([\bar{x}]_i))\|_2^2 \\ &\leq \frac{m}{D} - \frac{\|(f([x]_i) - f([x]_j)) + (f([\bar{x}]_j) - f([\bar{x}]_i))\|_2^2}{2} \\ &= \frac{m}{D} - \frac{\|(f([x]_i) - f([\bar{x}]_i) - (f([(x]_j) - f([\bar{x}]_j))\|_2^2}{2} \\ &\leq \frac{m}{D} - \frac{\frac{m}{D} + \frac{m}{D} - 2(f([x]_i) - f([\bar{x}]_i))(f([x]_j) - f([\bar{x}]_j))}{2} \\ &= (f([x]_i) - f([\bar{x}]_i))(f([x]_j) - f([\bar{x}]_j)). \end{aligned}$$

$\square$

Corollary 3.5. *For all $i, j \in [m]$ such that $|i-j| \leq \frac{m}{8D}$, and for any non-expansive map $f : V(H_m) \to \mathsf{NEG}$ with distortion at most D, the following inequality holds,*

$$(f([x]_i) - f([\bar{x}]_i))\,(f([x]_j) - f([\bar{x}]_j)) \geq \frac{m}{2D}.$$

Lemma 3.6. *Let $f : H_m \to \mathsf{NEG}$ be a non-expansive mapping. Furthermore, suppose that $\mathsf{dist}(f) \leq D$ and f is ε-efficient with respect to μ_m, then there exist an index $p \in [m]$ such that,*

$$\mathbb{E}_{\vec{e} \in \vec{E}(Q_m)} \mathbb{E}_{i,j \in \{p,\cdots,p+\ell\}} f([\vec{e}]_i) f([\vec{e}]_j) \lesssim \frac{D}{m} + \frac{\varepsilon m}{D}, \qquad (14)$$

where $\ell = \lfloor \frac{m}{8D} \rfloor$.

Proof. To prove this lemma first we bound the slack on sub-paths. Let s_p denote the expectation of the following quantum when $x \in Q_m$ and $k \in [m]$ are chosen uniformly at random

$$\sum_{r=p}^{p+\ell} \left(\|f([x]_r) - f([x \oplus e_k]_r)\|_2^2 + \|f([x]_{r+1}) - f([x \oplus e_k]_r)\|_2^2 \right) - \|f([x]_p) - f([x \oplus e_k]_{p+\ell})\|_2^2 .$$

Since f is ε-efficient with respect to μ_m, we have

$$\mathbb{E}_{p \in [m-\ell]} s_p \lesssim \varepsilon \ell,$$

and therefore there must exist an index p such that $s_p \lesssim \varepsilon \ell$. We show that any such p satisfies (14).

We bound the sum,

$$\mathbb{E}_{\vec{e} \in \vec{E}(Q_m)} \Big\| \sum_{i=p}^{p+\ell} f([\vec{e}]_i) \Big\|_2^2 \qquad (15)$$
$$\asymp \ell^2 \left(\mathbb{E}_{\vec{e} \in \vec{E}(Q_m)} \mathbb{E}_{i,j \in \{p,\cdots,p+\ell\}} f([\vec{e}]_i) f([\vec{e}]_j) \right)$$

by splitting it into two parts:

$$\Big\| \sum_{i=p}^{p+\ell} f([\vec{e}]_i) \Big\|_2^2 = \sum_{i=p}^{p+\ell} \|f([\vec{e}]_i)\|_2^2 + \sum_{i,j=p:i\neq j}^{p+\ell} f([\vec{e}]_i) f([\vec{e}]_j).$$

The map f is non-expansive, therefore $\|f([\vec{e}]_i)\|_2^2 \leq 1$, and

$$\Big\| \sum_{i=p}^{p+\ell} f([\vec{e}]_i) \Big\|_2^2 \lesssim \ell + \sum_{i,j=p:i\neq j}^{p+\ell} f([\vec{e}]_i) f([\vec{e}]_j).$$

By bounding,

$$\sum_{i,j=p:i\neq j}^{p+\ell} f([\vec{e}]_i) f([\vec{e}]_j), \qquad (16)$$

we can bound the overall sum and complete the proof.

For a given p and e, let

$$\nu_{p,\vec{e}} = \max_{i \neq j \in \{p,\ldots,p+\ell\}} f([\vec{e}]_i) f([\vec{e}]_j),$$

and let $i_0 < j_0$ be the indices were the maximum is achieved. We show that $s_p \lesssim \mathbb{E}_{e \in E(Q_m)} \nu_{p,\vec{e}}$ and then bound (15) by

$$\Big\| \sum_{i=p}^{p+\ell} f([\vec{e}]_i) \Big\|_2^2 \lesssim \ell + s_p \ell^2 \lesssim \ell^2 \left(\frac{1}{\ell} + \varepsilon \ell \right) \asymp \ell^2 \left(\frac{D}{m} + \frac{\varepsilon m}{D} \right),$$

to complete the proof.

For an edge $\vec{e} = (x, x \oplus e_k)$, we can write the triangle inequality for $\|\cdot\|_2^2$ among $[x]_{i_0}$, $[x]_{j_0}$ and $[x \oplus e_k]_{j_0}$ as

$$\begin{aligned}
0 &\geq (f([x]_{i_0}) - f([x]_{j_0})) \cdot (f([x]_{j_0}) - f([x \oplus e_k]_{j_0})) \\
&= (f([x]_{i_0}) - f([x \oplus e_k]_{i_0}) \cdot (f([x]_{j_0}) - f([x \oplus e_k]_{j_0}) \\
&\quad + f([x \oplus e_k]_{i_0}) - f([x]_{j_0})) \cdot (f([x]_{j_0}) - f([x \oplus e_k]_{j_0})) \\
&\geq \nu_{p,\vec{e}} + (f([x \oplus e_k]_{i_0}) - f([x]_{j_0})) \cdot (f([x]_{j_0}) - f([x \oplus e_k]_{j_0})),
\end{aligned}$$

hence

$$\|f([x \oplus e_k]_{i_0}) - f([x]_{j_0})\|_2^2 + \|f([x]_{j_0}) - f([x \oplus e_k]_{j_0})\|_2^2 \geq \|f([x \oplus e_k]_{i_0}) - f([x \oplus e_k]_{j_0})\|_2^2 + 2\nu_{p,\vec{e}}.$$

Using the triangle inequality, we obtain

$$\begin{aligned}
&\|f([x]_p) - f([x \oplus e_k]_{p+\ell})\|_2^2 \\
&\quad \leq \|f([x]_p) - f([x \oplus e_k]_{i_0})\|_2^2 \\
&\qquad + \|f([x \oplus e_k]_{i_0}) - f([x \oplus e_k]_{j_0})\|_2^2 \\
&\qquad + \|f([x \oplus e_k]_{j_0}) - f([x \oplus e_k]_{p+\ell})\|_2^2 \\
&\quad \leq -2\nu_{p,\vec{e}} + \sum_{r=p}^{p+\ell} \|f([x]_r) - f([x \oplus e_k]_r)\|_2^2 \\
&\qquad\qquad + \|f([x]_{r+1}) - f([x \oplus e_k]_r)\|_2^2 .
\end{aligned}$$

Therefore, $s_p \gtrsim \mathbb{E}_{\vec{e} \in \vec{E}(Q_m)} \nu_{p,\vec{e}}$ and the proof is complete. □

Now we can present the proof of Lemma 3.2.

Proof of Lemma 3.2. Let $f : H_m \to \mathsf{NEG}$ be a non-expansive map which is ε-efficient with respect to μ_m, where $\varepsilon = O(\frac{1}{m \log m})$. Let p satisfy the conclusion of Lemma 3.6. Let $f_i(x) = f([x]_i)$. Furthermore, let

$$\mathcal{F} = \left\{ f_i : p \leq i \leq p + \left\lfloor \frac{m}{8D} \right\rfloor \right\}.$$

An immediate application of Lemma 3.6 yields,

$$\mathbb{E}_{f,g \in \mathcal{F}, \vec{e} \in \vec{E}(Q_m)} [f(\vec{e}) \cdot g(\vec{e})] \lesssim \frac{D}{m} + \frac{\varepsilon m}{D}.$$

From Corollary 3.5 we have,

$$\mathbb{E}_{f,g \in \mathcal{F}, x \in V(Q_m)} [(f(x) - f(\bar{x})) \cdot (g(x) - g(\bar{x}))] \geq \frac{m}{2D}.$$

Applying Lemma 3.3 with the given bounds, we must have

$$m \gtrsim \frac{\frac{m}{D}}{\frac{D}{m} + \frac{\varepsilon m}{D}} \gtrsim \frac{1}{\varepsilon + \frac{D^2}{m^2}},$$

therefore

$$m\varepsilon + \frac{D^2}{m} \gtrsim 1,$$

and $D \gtrsim m^{1/2}$. □

Acknowledgements

We are grateful to Alex Jaffe for his participation in the early phases of this work. We also thank Sushant Sachdeva for pointing out a quantitative improvement in our lower bound argument.

4. REFERENCES

[ACMM05] A. Agarwal, M. Charikar, K. Makarychev, and Y. Makarychev. $O(\sqrt{\log n})$ approximation algorithms for Min UnCut, Min 2CNF Deletion, and directed cut problems. In *37th Annual ACM Symposium on Theory of Computing*. ACM, 2005.

[AHK04] S. Arora, E. Hazan, and S. Kale. $O(\sqrt{\log n})$ approximation to SPARSEST CUT in $\tilde{O}(n^2)$ time. In *45th Annual Syposium on Foundations of Computer Science*, pages 238–247. IEEE Computer Society, 2004.

[ALN08] Sanjeev Arora, James R. Lee, and Assaf Naor. Euclidean distortion and the Sparsest Cut. *J. Amer. Math. Soc.*, 21(1):1–21, 2008.

[AR98] Yonatan Aumann and Yuval Rabani. An $O(\log k)$ approximate min-cut max-flow theorem and approximation algorithm. *SIAM J. Comput.*, 27(1):291–301 (electronic), 1998.

[ARV04] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings, and graph partitionings. In *36th Annual Symposium on the Theory of Computing*, pages 222–231, 2004.

[Ass83] Patrice Assouad. Plongements lipschitziens dans $\mathbf{R}^n$. *Bull. Soc. Math. France*, 111(4):429–448, 1983.

[BL00] Yoav Benyamini and Joram Lindenstrauss. *Geometric nonlinear functional analysis. Vol. 1*, volume 48 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2000.

[Bou85] J. Bourgain. On Lipschitz embedding of finite metric spaces in Hilbert space. *Israel J. Math.*, 52(1-2):46–52, 1985.

[CGR05] S. Chawla, A. Gupta, and H. Räcke. An improved approximation to sparsest cut. In *Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms*, Vancouver, 2005. ACM.

[Che99] J. Cheeger. Differentiability of Lipschitz functions on metric measure spaces. *Geom. Funct. Anal.*, 9(3):428–517, 1999.

[CK06a] J. Cheeger and B. Kleiner. Differentiating maps into L^1 and the geometry of BV functions. arXiv:math.MG/0611954, 2006.

[CK06b] Jeff Cheeger and Bruce Kleiner. Generalized differential and bi-Lipschitz nonembedding in L^1. *C. R. Math. Acad. Sci. Paris*, 343(5):297–301, 2006.

[CK06c] Jeff Cheeger and Bruce Kleiner. On the differentiability of Lipschitz maps from metric measure spaces to Banach spaces. In *Inspired by S. S. Chern*, volume 11 of *Nankai Tracts Math.*, pages 129–152. World Sci. Publ., Hackensack, NJ, 2006.

[CK09] J. Cheeger and B. Kleiner. Metric differentiation, monotonicity and maps to L^1. arXiv:0907.3295, 2009.

[CKN09] J. Cheeger, B. Kleiner, and A. Naor. A $(\log n)^{\Omega(1)}$ integrality gap for the Sparsest Cut SDP. In *50th Annual Symposium on Foundations of Computer Science*. IEEE Computer Soc., Los Alamitos, CA, 2009.

[CMM06] Eden Chlamtac, Konstantin Makayrchev, and Yury Makarychev. How to play unique game using embeddings. In *47th Annual Syposium on Foundations of Computer Science*, 2006.

[DKSV06] N. Devanur, S. Khot, R. Saket, and N. K. Vishnoi. Integrality gaps for sparsest cut and minimum linear arrangement problems. In *38th Annual Symposium on the Theory of Computing*, 2006.

[EFW06] A. Eskin, D. Fisher, and K. Whyte. Quasi-isometries and rigidity of solvable groups. Preprint, 2006.

[Enf69] P. Enflo. On the nonexistence of uniform homeomorphisms between L_p-spaces. *Ark. Mat.*, 8:103–105, 1969.

[FHL05] U. Feige, M. T. Hajiaghayi, and J. R. Lee. Improved approximation algorithms for minimum-weight vertex separators. In *37th Annual ACM Symposium on Theory of Computing*. ACM, 2005. To appear, *SIAM J. Comput.*

[GNRS99] A. Gupta, I. Newman, Y. Rabinovich, and A. Sinclair. Cuts, trees and l_1-embeddings of graphs. In *40th Annual Symposium on Foundations of Computer Science*, pages 399–408. IEEE Computer Soc., Los Alamitos, CA, 1999.

[Kar09] George Karakostas. A better approximation ratio for the vertex cover problem. *ACM Trans. Algorithms*, 5(4):Art. 41, 8, 2009.

[KL08] A. Kolla and J. R. Lee. Sparsest cut on quotients of the hypercube. Preprint, 2008.

[KLMN05] R. Krauthgamer, J. R. Lee, M. Mendel, and A. Naor. Measured descent: a new embedding method for finite metrics. *Geom. Funct. Anal.*, 15(4):839–858, 2005.

[KR06] Robert Krauthgamer and Yuval Rabani. Improved lower bounds for embeddings into l_1. In *Proceedings of the 17th annual ACM-SIAM symposium on Discrete algorithm*, pages 1010–1017. ACM Press, 2006.

[KV05] Subhash Khot and Nisheeth 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*, pages 53–62, 2005.

[Lee05] James R. Lee. On distance scales, embeddings, and efficient relaxations of the cut cone. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 92–101, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[Lee09] J. R. Lee. Volume distortion for subsets of Euclidean spaces. *Discrete Comput. Geom.*, 41(4):590–615, 2009.

[Lin02] Nathan Linial. Finite metric-spaces—combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians, Vol. III (Beijing, 2002)*, pages 573–586, Beijing, 2002. Higher Ed. Press.

[LLR95] N. Linial, E. London, and Y. Rabinovich. The geometry of graphs and some of its algorithmic applications. *Combinatorica*, 15(2):215–245, 1995.

[LMN05] James R. Lee, Manor Mendel, and Assaf Naor. Metric structures in L_1: dimension, snowflakes, and average distortion. *European J. Combin.*, 26(8):1180–1190, 2005.

[LN06] J. R. Lee and A. Naor. L_p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *47th Annual Symposium on Foundations of Computer Science*. IEEE Computer Soc., Los Alamitos, CA, 2006.

[LR99] Tom Leighton and Satish Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *J. ACM*, 46(6):787–832, 1999.

[LR07] J. R. Lee and P. Rhagevendra. Coarse differentiation and multi-flows in planar graphs. To appear, *Disc. Comp. Geom.*, 2007.

[LS09] J. R. Lee and A. Sidiropoulos. On the geometry of graphs with a forbidden minor. In *STOC'09: Proceedings of the 41st Annual ACM Symposium on Theory of Computing*. ACM, 2009.

[Mat02a] J. Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[Mat02b] Jiri Matoušek. Open problems on embeddings of finite metric spaces. *Available at http://kam.mff.cuni.cz/∼matousek/haifaop.ps*, 2002.

[Pan89] Pierre Pansu. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Ann. of Math. (2)*, 129(1):1–60, 1989.

[Rao99] Satish 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.

[She09] J. Sherman. Breaking the multicommodity flow barrier for $\sqrt{\log n}$-approximations to Sparsest Cut. In *50th Annual Symposium on Foundations of Computer Science*. IEEE Computer Soc., Los Alamitos, CA, 2009.