Near-optimal distortion bounds for embedding doubling spaces into L_1

[extended abstract]

James R. Lee   Anastasios Sidiropoulos

November 4, 2010

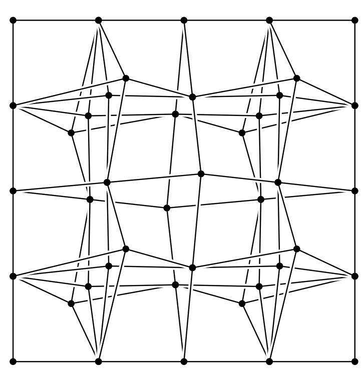
Abstract

We exhibit an infinite doubling metric space $(\mathcal{X}, d)$ such that for any non-expansive mapping $f : \mathcal{X} \to L_1$, there exists a pair $x, y \in \mathcal{X}$ with $d(x, y)$ arbitrarily large, and such that

$$\frac{\|f(x) - f(y)\|_1}{d(x, y)} \lesssim \sqrt{\frac{\log \log d(x, y)}{\log d(x, y)}}.$$

As a consequence, we show that there are n-point doubling metrics which require distortion $\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$ into L_1, matching the upper bound of [Gupta-Krauthgamer-Lee, FOCS'03] up to a factor of $O(\sqrt{\log \log n})$. The best previous lower bound for doubling spaces, due to [Cheeger-Kleiner-Naor, FOCS'09] was of the form $(\log n)^\delta$ for some small, unspecified value of $\delta > 0$.

Furthermore, this gives a nearly optimal integrality gap for a weak version of the SDP for the general Sparsest Cut Problem. The weak SDP suffices for all known rounding algorithms, and the best previous gap was of the order $\frac{(\log n)^{1/4}}{\log \log n}$ [Lee-Moharrami, STOC'10]. We conjecture that our construction admits an equivalent metric of negative type. Resolution of this conjecture would lead to an integrality gap of $\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$ for the Goemans-Linial SDP, nearly matching the upper bound $O(\sqrt{\log n} \log \log n)$ of [Arora-Lee-Naor, STOC'05].



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. 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 exists a mapping $f : X \to L_2$ such that

$$\|f(x) - f(y)\|_2^2 = d(x, y)$$

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 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 $c_1(X, d) \leq O(\sqrt{\log n} \log \log n)$ 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 basic optimization 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. On the other hand, since L_1 metrics correspond precisely to the cut cone (whose extreme points are exactly the cuts on a given set of points), understanding L_1-embeddability of families is of great importance in combinatorial optimization. The present paper makes substantial progress on both fronts, as we now discuss.

Integrality gaps and weak negative type. 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. Recall that a metric space is of negative type if $c_2(X, \sqrt{d}) = 1$. We will say that (X, d) is a space of D-weak negative type if $c_2(X, \sqrt{d}) \le D$. In all the above constructions of integrality gaps, it is relatively easy to show that the space in question is $O(1)$-weak negative type. 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 a space of $O(1)$-weak negative type. Indeed, in all these cases, this fact was taken as evidence and motivation that eventually a (strong) negative-type metric could be constructed.

Our result. We show that there exist arbitrarily large n-point metric spaces of $O(1)$-weak negative type that require distortion

$$\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$$

to embed into L_1. This almost matches the upper bound of $O(\sqrt{\log n} \log \log n)$ from [ALN08], which also holds for $O(1)$-weak negative type metrics. The best previous lower bound, due to [LM10], is on the order of $(\log n)^{1/4}$, up to a factor of $O(\log \log n)$. As we discuss below, our lower bound also yields a nearly-optimal integrality gap for an SDP which, while weaker than the Goemans-Linial SDP, is still capable of achieving the best-known approximation algorithms. Indeed, the full negative-type constraint has not been used in any rounding analysis that we are aware of; in all such algorithms, the $O(1)$-weak negative type constraint suffices. Furthermore, we conjecture that our space embeds into a space of negative type; this would imply a nearly-optimal integrality gap for the Goemans-Linial SDP.

A metric space (X, d) is called *doubling* if every ball in X can be covered by $O(1)$ balls of half the radius. Our proof shows more: There exists a doubling metric space (X, d) such that arbitrarily large n-point subsets of X require distortion $\Omega\left(\sqrt{\frac{\log n}{\log \log n}}\right)$ to embed in L_1. By Assouad's embedding theorem [Ass83], every doubling space is of $O(1)$-weak negative type. Our lower bound for doubling spaces nearly matches the $O(\sqrt{\log n})$ upper bound of [GKL03]. The best previous result of [CKN09] gives a lower bound of $(\log n)^{\delta_0}$ for a very small value of $\delta_0 > 0$. We remark that proving lower bounds into L_1 is significantly more difficult than for other L_p spaces. In fact, for every fixed $p > 1$, the best asymptotic distortion is $(\log n)^{\min(\frac{1}{2},\frac{1}{p})}$ [GKL03].

Indeed, proving lower bounds against embeddability into L_1 has been a notoriously difficult challenge. It was asked in [Mat02b] whether every "$O(1)$-decomposable" metric admits an $O(1)$-distortion embedding into L_1. In particular, every doubling metric is $O(1)$-decomposable [GKL03]. This was only recently answered negatively in [CK06a], resolving a conjecture from [LN06]. The main novelty of [CK06a] is to develop a differentiation theory for L_1-valued maps, and employ this in proving distortion lower bounds. The bound of [CK06a] is non-quantitative, and achieving a bound of the form $(\log n)^{\delta_0}$ for some $\delta_0 > 0$ required first a new approach to the qualitative non-embedding result [CK09], and then a very technical and difficult effort [CKN09] to obtain a concrete bound. It should be noted that the latter effort faced significant challenges because one must work at a "definite scale" instead of passing to a limit object. We face a similar challenge in the present work, except that in order to obtain the correct asymptotic dependence, many further obstacles arise. On the other hand, in many ways our construction is simpler than the Heisenberg group, and this allows us to present a lower bound which is nearly-optimal, and also more accessible than that of [CKN09]. The details of our approach are discussed subsequently. First, we address the issue of weak vs. strong negative type.

Weak vs. strong negative type. Unfortunately, it was shown in [LM10] that there are metric spaces of $O(1)$-weak negative type which cannot be embedded with $O(1)$ distortion into a space of genuine negative type. In fact, for n-point spaces, the gap between weak and strong negative type can be as bad as $\Omega\left(\frac{(\log n)^{1/4}}{\log\log n}\right)$. Nevertheless, these spaces are far from doubling. A central property used for the lower bound space X in [LM10] is that any Lipschitz mapping of X into L_2 must shrink the average diameter of X by an arbitrarily large amount. No such property can hold for doubling spaces. We conjecture that our lower bound space does, in fact, admit an equivalent metric of negative type.

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 $w_1, w_2, \ldots, w_k \in W$,

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

for some constant $C = O(1)$, independent of the sequence. This is merely the weak negative type condition in disguise: It simply says that W is the image of a weak negative type embedding of some metric space. In all known algorithmic applications, it is only the weak condition that is needed.

Far more than being a curiousity of the analysis, the fact that a weaker condition suffices is actually the basis for algorithms 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 fixed $\varepsilon > 0$. 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. As stated before, our lower bound yields a nearly-optimal integrality gap for these weaker programs, for the case of general Sparsest Cut.

Differentiation and bi-Lipschitz embeddings into L_1. 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 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, and work of [LM10] shows that it can also be used to prove lower bounds against embeddings into negative type metrics. In Section 2, we proceed to a detailed discussion of our construction and approach. For the moment, we say a few general words about the new obstacles we face in the present work.

At a very broad level, distortion lower bounds proved via differentiation proceed as follows. One first argues that any low-distortion embedding must be quantitatively very well-controlled on a small piece of the lower bound space. For instance, consider a model statement: Every Lipschitz mapping $f : \mathbb{R} \to \mathbb{R}$ is differentiable almost everywhere, i.e. is almost everywhere locally well-approximated by a linear function. In our case, the lower bound space is constructed so that the differentiation theory need only be applied to mappings $f : [0,1]^2 \to L_1$ from the unit square to L_1.

For L_1-valued mappings, our proof follows the differentiation approach developed in [CK09] and [LR07], where the local conclusion is that the cut decomposition of the embedding must have most of its weight concentrated on cuts that are "monotone" with respect to a given family of lines. The next step is to prove a structural theorem which classifies the structure of nearly monotone sets. Here is the first place at which our approach must deviate substantially from previous ones. Since we are shooting for a precise quantitative dependence, we cannot say that almost all the sets are "nearly monotone," only that their average monotonicity is small. Thus we must classify the structure, not only of 99%-structured sets, but also of 1%-structured sets.

The second major difficulty is that since we must work at a fixed scale, we do not have monotonicity with respect to lines, but instead with respect to discrete sequences of points along these lines. If these points were deterministic, then even very bad sets could elude our test lines by being periodic in sync with the discrete sequence. To counter this, we develop a "random" discrete differentiation theory. For instance, consider a mapping $f : [0,1] \to L_1$. Instead of subdividing $[0,1]$ into a sequence of hierarchical partitions, we use a sequence of random progressively finer, *non-hierarchical* subdivisions. We believe this technique will be useful in future quantitatively optimal bounds for L_1 embeddings.

2 Overview

First, we describe our constructions. Then in Section 2.1, we give a qualitative proof that our lower bound space does not embed into L_1 with $O(1)$ distortion. Finally, in Section 2.2, we describe the

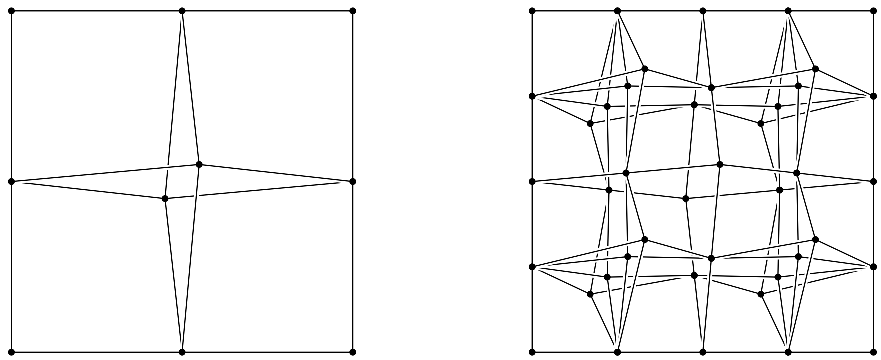


Figure 1: The spaces $\mathcal{X}_1$ and $\mathcal{X}_2$ after one and two rounds of gluing, respectively.

novel aspects that go into proving a precise quantitative bound. Pointers are given to the appendix, where the formal arguments appear.

The diamondfold and the Laakso-fold. Informally, our construction can be defined inductively as follows. We start with $\mathcal{X}_0$ which is just a copy of $[0,1]^2$. The space $\mathcal{X}_1$ is two copies of $\mathcal{X}_0$ glued together along $\partial[0,1]^2$, where we use ∂S to denote the topological boundary of a set S. In general, we will take the metric on such a gluing as the quotient metric on $([0,1]^2 \amalg [0,1]^2)/\sim$, where $\amalg$ denotes the disjoint union, and $\sim$ is the equivalence relation along the boundary. See Figure 1, which contains a topological representation of $\mathcal{X}_1$. (In the geometry of $\mathcal{X}_1$, both copies of $[0,1]^2$ are flat.)

Now, in $\mathcal{X}_1$ there are 8 natural subsquares present in Figure 1. The space $\mathcal{X}_2$ arises after applying the same gluing process to each of these 8 subsquares. The space $\mathcal{X}_3$ arises from the gluing process applied to each of the 64 subsquares in $\mathcal{X}_2$, and so on. We refer to the spaces $\{\mathcal{X}_k\}$ as the *diamondfolds,* named after the diamond graphs of Newman and Rabinovich [NR03]. While it is possible to pass to a Gromov-Hausdorff limit of these spaces (a limit space contains a copy of every $\mathcal{X}_k$ isometrically), we defer such a discussion from the present abstract, as it is non-essential in proving our lower bound.

The diamondfolds are *not* doubling, but they are slightly easier to reason about than their doubling counterparts, the *Laakso-folds,* which we denote by $\{\mathcal{L}_k\}$. The Laakso-folds are based on the Laakso graphs [Laa02, LP01, GKL03]. These are constructed in the same manner as the diamondfolds, except the gluing process occurs only along a Cantor set. We refer to Figure 2 for a graphical description. The precise definitions of both spaces occur in Section A.1.

We remark that if the reader wants to think about graphs instead of continuous spaces, then it is comforting to know that our distortion lower bound will actually hold for the graph formed by taking the 1-dimensional simplicial complex which is composed of the boundaries of all the glued squares. However, this is only because the vertices of this complex converge to a net in some other $\mathcal{X}_k$ or $\mathcal{L}_k$ space. Our proof requires the full continuous ambient space, and the discretization will be done by a Lipschitz extension argument.

A key property. We now state a key property of the diamondfold construction. A similar property holds in the Laakso-fold case. Observe that the gluing $([0,1]^2 \amalg [0,1]^2)/\sim$ is topologically a sphere. Call any such (possibly scaled) sphere in the construction of some $\mathcal{X}_k$ space an *identified sphere.* Clearly every identified sphere has a "top" sheet and a "bottom" sheet. These identified spheres

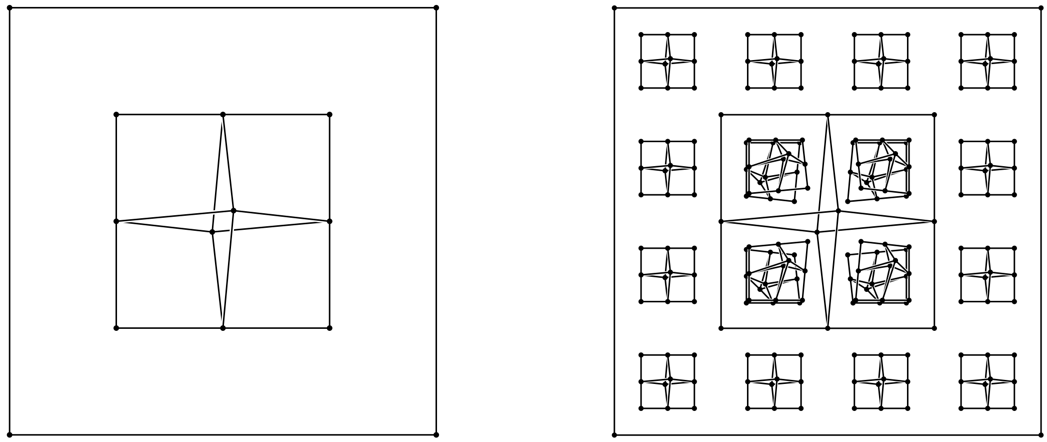


Figure 2: The Laakso-folds after one and two rounds of gluing, respectively.

also have *levels* corresponding to the round of the construction in which they were formed. For concreteness, note that $\mathcal{X}_2$ has 9 identified spheres: One level-1 sphere, and 8 level-2 spheres.

Corresponding to these spheres, we can consider a dyadic partitioning of $[0,1]^2$. For $k = 0, 1, \ldots$ let $\mathcal{D}_k$ be the set of closed squares in $[0,1]^2$ whose corners occur at multiples of 2^{-k}. In a natural way, the set of level-k spheres correspond to the set $\mathcal{D}_k$. The next property is straightforward, but will be crucial.

Proposition 2.1 (Monochromatic sheet)**.** *Suppose that following holds for some level $j \leq k$: Every level-j identified sphere in $\mathcal{X}_k$ has either its top sheet or bottom sheet colored red. Then there is an isometric copy of $[0,1]^2$ contained in $\mathcal{X}_k$ such that every dyadic square in $\mathcal{D}_j$ is colored red.*

2.1 A qualitative lower bound

We first give an intuitive (and ultimately invalid) sketch of why $c_1(\mathcal{X}_k) \to \infty$. To do so, we need to discuss L_1-valued mappings and their cut cone representation. We will ignore issues of measurability for this informal description; we remark only that they play a very mild role in the formal arguments. Let X be a set, and consider a mapping $f : X \to L_1$. It is well-known that there exists a measure ν on 2^X, the subsets of X, such that for any $x, y \in X$,

$$\|f(x) - f(y)\|_1 = \int |\mathbf{1}_S(x) - \mathbf{1}_S(y)| \, d\nu(S).$$

Now, examine a Lipschitz mapping $f : [0,1]^2 \to L_1$, and let ν be the corresponding cut measure. Let $\ell \subseteq \mathbb{R}^2$ be some line for which $\ell \cap [0,1]^2 \neq \emptyset$. The point is now that we can say a lot about the local structure of the map $f|_\ell : \ell \to L_1$ and the corresponding cut measure ν_ℓ on ℓ: By a differentiation argument, one can show, morally, that on small subintervals $[a, b] \subseteq \ell$, the cut measure ν_ℓ restricted to $[a, b]$ is concentred almost entirely on half-segments of the form $[a, b']$ for some $b' \leq b$. This "montoncity" property was observed independently in [CK09] (who use the metric differentiation theory of Pauls [Pau01]) and in [LR07] (where the coarse differentiation of Eskin, Fisher, and Whyte [EFW06] is employed).

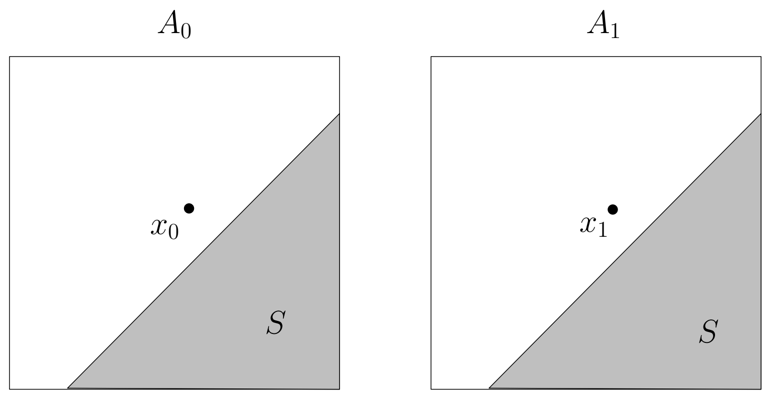


Figure 3: The top sheet and bottom sheet of the sphere $\mathbb{S}$. A_0 and A_1 have a common boundary.

When this information is combined from all lines ℓ, and a proper union bound is taken into account, one recovers (again, only morally) the following: For a sufficiently large value of k, for most squares $A \in \mathcal{D}_k$, the following holds: If ν_A is the cut measure restricted to A, then ν_A is concentrated almost entirely on cuts which are formed by intersecting a halfspace with A, e.g. see Figure 3.

Let us assume, for the moment, a stronger conclusion: That for every such mapping $f : [0,1]^2 \to L_1$, there exists a square $A \in \mathcal{D}_k$, where the restricted cut measure ν_A is concentrated *entirely* on halfspace cuts. In this case, given a 1-Lipschitz mapping $f : \mathcal{X}_k \to L_1$, we know that there must exist a level-k identified sphere $\mathbb{S}$ in $\mathcal{X}_k$, such that if A_0 is the top sheet of $\mathbb{S}$ and A_1 is the bottom sheet, then the restricted cut measures ν_{A_0} and ν_{A_1} are both concentrated entirely on halfspace cuts. Otherwise, we could apply Proposition 2.1 and conclude that there is an isometric copy of $[0,1]^2$ where all $A \in \mathcal{D}_k$ have some of their cut measure *not on* halfspace cuts, and this would contradict our assumption.

Now, take these sheets A_0 and A_1, and let $\nu_{\mathbb{S}}$ be the cut measure restricted to $\mathbb{S}$. The main point is now this: For any cut $S \subseteq \mathbb{S}$ such that $A_0 \cap S$ and $A_1 \cap S$ are both halfspaces, they must be the *same* halfspace of A_0 and A_1 (under the canonical identification). This is because a halfspace cut $S \subseteq A_0$ is completely determined by the intersection $S \cap \partial A_0$, and $\partial A_0 = \partial A_1$ since A_0 and A_1 are glued together along their boundaries. Thus for purely "vertically separated" points $x_0 \in A_0$ and $x_1 \in A_1$ as in Figure 3, i.e. points which are equal under the canonical identifications of A_0 and A_1, it is impossible to have $\mathbf{1}_S(x_0) \neq \mathbf{1}_S(x_1)$. Thus under our assumption, the $\nu_{\mathbb{S}}$-measure of cuts which separate x_0 and x_1 is 0. In other words, $f(x_0) = f(x_1)$. Since $d_{\mathcal{X}_k}(x_0, x_1) > 0$, this implies that f has infinite distortion, completing our qualitative sketch.

Of course, this sketch is not mathematically valid (in particular, it is possible to show that $c_1(\mathcal{X}_k) \lesssim \sqrt{k}$), but it gives some structure and intuition to the quantitative arguments to come. We remark that this is not merely a "brute force" quantification (e.g. "chasing ε's and δ's"). Even formalizing the above argument is highly non-trivial. Furthermore, obtaining any explicit bound already requires a difficult classification argument. Finally, obtaining an bound which is asymptotically near-optimal requires all ingredients to fit together seamlessly, with essentially no loss at any step.

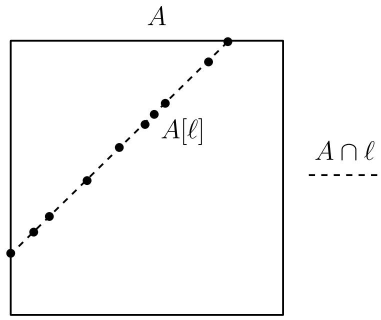


Figure 4: The square A, the line segment $A \cap \ell$, and the sprinkled points $A[\ell]$.

2.2 The quantitative lower bound

Making the preceding argument formal requires two major ingredients. The first is a differentiation theorem that gives the right kind of control on the cut measure restricted to a small identified sphere. The second is a classification of the "controlled" sets. Furthermore, these two ingredients have to make a trade-off; the classification theorem wants as much as possible from the differentiation step, but we cannot ask for too much without losing in the distortion lower bound (because the differentiation argument will require more scales, increasing the number of points).

Setup: The kinematic measure and randomly sprinkled needles. Let μ_0 be the kinetmatic measure on lines $\ell \subseteq \mathbb{R}^2$. This is the unique measure, up to scaling, which is invariant under rigid motions. Let Λ be the set of all lines in $\mathbb{R}^2$ that intersect $[0,1]^2$, and define the measure $\mu(S) = \mu_0(S \cap \Lambda)$, scaled so that $\mu(\Lambda) = 1$.

Let $\mathcal{D} = \bigcup_{j \geq 0} \mathcal{D}_j$ be the set of *all* dyadic squares in $[0,1]^2$. Fix some parameter $k \in \mathbb{N}$. For any dyadic square $A \in \mathcal{D}$, let $A[\ell]$ denote a random subset described as follows: $A[\ell] = (\ell \cap \partial A) \cup T_{A,\ell}$, where $T_{A,\ell}$ is a uniformly random subset of $A \cap \ell$ of size k. The random sets $T_{A,\ell}$ are taken to be independent of each other. We will use $\mathbb{E}(\cdot)$ to denote expectation over these random sets.

The differentiation step. We introduce some notation. Let (Y, d_Y) be an arbitrary meric space. For a function $F : [0,1]^2 \to Y$ and any finite sequence $s = \langle s_1, s_2, \ldots, s_j \rangle$ of points in $[0,1]^2$, define

$$\Delta_F(s) = \sum_{i=1}^{j-1} d_Y(F(s_i), F(s_{i+1})),$$

i.e. the variation of F along s. The following result is proved in Section B.

Theorem 2.2. *For every metric space Y and 1-Lipschitz mapping $F : [0,1]^2 \to Y$, and every $0 < \varepsilon < 1$, there exists a dyadic square $A \in \mathcal{D}$ of side length at least 2^{-r}, with $r \lesssim \frac{1}{\varepsilon} \log \frac{k}{\varepsilon}$, and such that*

$$\int \mathbb{E}\left[\Delta_F(A[\ell])\right] d\mu(\ell) \leq \varepsilon 2^{-2r} + \int \Delta_F(\partial A \cap \ell)\, d\mu(\ell). \tag{1}$$

Before interpreting this theorem, let us scale is up to assume that $A = [0,1]^2$, in which case it says that

$$\int \mathbb{E}\left[\Delta_F(A[\ell])\right] d\mu(\ell) \leq O(\varepsilon) + \int \Delta_F(\partial A \cap \ell)\, d\mu(\ell). \tag{2}$$

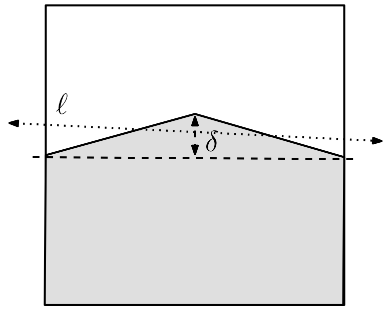


Figure 5: A cut with symmetric difference $\Omega(\delta)$ to a halfplane, but $\theta_S^k = O(\delta^2)$.

(Here, we have used the fact that the μ-measure of lines which intersect a square of side length at least 2^{-r} is $\gtrsim 2^{-2r}$.)

This says that if we choose a random line ℓ and a random set $A[\ell]$, then the expected variation of F over the the whole set $A[\ell]$ is only $O(\varepsilon)$ more than the variation of F over the endpoints $\partial A \cap \ell$. (See Figure 4.)

If we now assume that $F : [0,1]^2 \to L_1$, and let ν be the corresponding cut measure on A, then after an application of Fubini's theorem, the conclusion is that

$$\int \int \mathbb{E}\left[\Delta_{\mathbf{1}_S}(A[\ell]) - \Delta_{\mathbf{1}_S}(\partial A \cap \ell)\right] d\mu(\ell)\, d\nu(S) \le O(\varepsilon). \tag{3}$$

For a set S, let

$$\theta_S^k = \int \mathbb{E}\left[\Delta_{\mathbf{1}_S}(A[\ell]) - \Delta_{\mathbf{1}_S}(\partial A \cap \ell)\right] d\mu(\ell),$$

where we recall that k is the number of random points composing $A[\ell]$. Observe that $\Delta_{\mathbf{1}_S}(A[\ell])$ represents the number of times that we cross S when walking along the points of $A[\ell]$, while $\Delta_{\mathbf{1}_S}(\partial A \cap \ell)$ is 0 if ℓ crosses S and 1 otherwise. The important observation is that these two quantities are precisely equal when A is the intersection of S with a halfspace. In other words, $\theta_{H\cap A}^k = 0$ for a halfspace H. We know from (3) that $\int \theta_S^k \, d\nu(S) \le O(\varepsilon)$. Hence our next goal, the "classification" step, is to show the reverse implication: if θ_S is small, then S is close to a halfspace cut.

One comment before we move on: We are measuring the variation of F with respect to a random subdivision in Theorem 2.2. A more standard approach would measure variation along a fixed sequence of partitions of $[0,1]^2$, each a refinement of the previous. It is important for the classification step that points of $A[\ell]$ are sometimes allowed to be arbitrarily close together. Our observation is simply that as long as the random subdivisions become increasingly dense, they will behave essentially like a sequence of "lazy" refinements. The error from anomalous events is easily controlled using the Lipschitz property of F.

The classification step. Our next goal is to prove results of the following sort (see Lemma C.8). We use λ_2 to denote the Lebesgue measure on $\mathbb{R}^2$.

Lemma 2.3 (Classification Lemma). *For every $k \in \mathbb{N}$, the following holds. Let $S \subseteq [0,1]^2$ be a measurable set. Then there exists a halfspace H such that*

$$\lambda_2(S \triangle H) \lesssim \sqrt{\theta_S^k} + k^{-1/12}.$$

In other words, sets with θ_S^k small are close to halfspaces in symmetric difference, up to the coarseness of our observations (controlled by the parameter k). We remark that the $\sqrt{\cdot}$ dependence is tight, and is the ultimate source of power $\frac{1}{2}$ in our $\sqrt{\frac{\log n}{\log\log n}}$ bound. To see this, consider the set in Figure 5. It has symmetric difference $\Omega(\delta)$ to a halfspace, yet the measure of lines that intersect the boundary twice is only $O(\delta^2)$. The main point of Lemma 2.3 is as a tool in proving the following bound on pairs of sets. (See Lemma C.11.)

Theorem 2.4. *For every $k \in \mathbb{N}$, the following holds. Let $S, S' \subseteq [0,1]^2$ be measurable sets such that $S \cap \partial[0,1]^2 = S' \cap \partial[0,1]^2$. Then,*

$$\lambda_2(S \triangle S') \lesssim \sqrt{\theta_S^k + \theta_{S'}^k} + k^{-1/12}.$$

To finish the argument, we need one additional theorem which relies on the classification lemma. (See Lemma C.12.)

Theorem 2.5. *For every $k \in \mathbb{N}$, the following holds. If $S \subseteq [0,1]^2$ satisfies $\lambda_2(S) \le 1/64$, then*

$$\lambda_2\left(S \cap [\tfrac{1}{4}, \tfrac{3}{4}]^2\right) \lesssim \theta_S^k + k^{-1/6}.$$

Putting everything together. We are now ready to prove the distortion lower bound. Let $\varepsilon > 0$ be given, let $k = \lceil \varepsilon^{-12} \rceil$, and choose an integer $r \lesssim \frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$ so that the conclusion of Theorem 2.2 follows. Let $N \subseteq \mathcal{X}_r$ be a δ-net, for some δ to be specified later. Note that $|N| \le O(2^r \delta^{-2})$.

Suppose that $F_0 : N_\delta \to L_1$ is a 1-Lipschitz mapping with distortion D. A simple Lipschitz extension argument shows that there is an $O(1)$-Lipschitz extension $F : \mathcal{X}_r \to L_1$. This can be shown by hand, but for instance it is an easy observation that the space $\mathcal{X}_r$ is $O(1)$-decomposable in the sense of [LN04], and thus the required extension exists. (In fact, it is easy to see that the extension can be taken so that the cut measure induced by F is supported on closed, polygonal sets in $\mathcal{X}_k$.) By rescaling F_0, we may assume that F is 1-Lipschitz, and F_0 has distortion $O(D)$.

By choosing $\delta \ge 2^{-r}/O(D)$, we guarantee that for any induced sphere $\mathbb{S}$ in $\mathcal{X}_r$, if A_0 and A_1 are the top and bottom sheets of $\mathbb{S}$, then

$$\int_{x \in [\frac{1}{4}, \frac{3}{4}]^2} \|F(\pi_0(x)) - F(\pi_1(x))\|_1 \, dx \gtrsim \frac{\mathrm{diam}(A_0)}{D}, \tag{4}$$

where π_0, π_1 are the natural maps from $[0,1]^2$ to A_0 and A_1, respectively. This follows from $d_{\mathcal{X}_r}(\pi_0(x), \pi_1(x)) \gtrsim \mathrm{diam}(A_0)$ for $x \in [\frac{1}{4}, \frac{3}{4}]^2$.

Applying Theorem 2.2 and the monochromatic sheet principle (Proposition 2.1), we find an induced sphere $\mathbb{S}$ in $\mathcal{X}_r$ such that the conclusion of Theorem 2.2 holds for both the top and bottom sheets, A_0 and A_1, of $\mathbb{S}$. Scaling so that A_0 and A_1 are both isometric to $[0,1]^2$, and letting ν be the cut measure induced by F, restricted to $\mathbb{S}$, we have

$$\int \theta_{S \cap A_0}^k + \theta_{S \cap A_1}^k \, d\nu(S) \lesssim \varepsilon. \tag{5}$$

On the other hand, after scaling, (4) gives

$$\int \int_{x \in [\frac{1}{4}, \frac{3}{4}]^2} |\mathbf{1}_S(\pi_0(x)) - \mathbf{1}_S(\pi_1(x))| \, dx \, d\nu(S) \gtrsim \frac{1}{D}.$$

We now write ν_{inner} for the measure ν restricted to sets S which satisfy

$$(\pi_0^{-1}(S) \cup \pi_1^{-1}(S)) \cap [\tfrac{1}{4}, \tfrac{3}{4}]^2 \neq \emptyset.$$

Furthermore, we decompose this measure into three disjointly supported measures

$$\nu_{\text{inner}} = \nu_{\text{large}} + \nu_{\text{small}} + \nu_{\text{tiny}},$$

where the three measures are supported on sets for which: One of the sets $S \cap A_0$ or $S \cap A_1$ has measure at least $1/64$ (ν_{large}), both sets have measure at most $1/64$, but one set has measure greater than ε (ν_{small}), both sets have measure at most ε.

Now, write $\rho(S) = \lambda_2(\pi_0^{-1}(A_0 \cap S) \triangle \pi_1^{-1}(A_1 \cap S))$, and then

$$\begin{aligned}
\frac{1}{D} &\lesssim \int \int_{x \in [\frac{1}{4}, \frac{3}{4}]^2} |\mathbf{1}_S(\pi_0(x)) - \mathbf{1}_S(\pi_1(x))| \, dx \, d\nu_{\text{inner}}(S) \\
&\leq 2 \int \lambda_2(\pi_0^{-1}(A_0 \cap S) \triangle \pi_1^{-1}(A_1 \cap S)) \, d\nu_{\text{inner}}(S) \\
&\lesssim \int \rho \, d\nu_{\text{large}} + \int \rho \, d\nu_{\text{small}} + \int \rho \, d\nu_{\text{tiny}}.
\end{aligned}$$

We bound each of these three terms separately.

First, by Theorem 2.4, and using Cauchy-Schwarz, we have

$$\begin{aligned}
\int \rho \, d\nu_{\text{large}} &\lesssim \int \sqrt{\theta^k_{A_0 \cap S}} + \sqrt{\theta^k_{A_1 \cap S}} + k^{-1/12} \, d\nu_{\text{large}} \\
&\lesssim \sqrt{\int d\nu_{\text{large}}} \sqrt{\int \theta^k_{A_0 \cap S} + \theta^k_{A_1 \cap S} \, d\nu_{\text{large}}} + \int k^{-1/12} \, d\nu_{\text{large}} \\
&\lesssim \sqrt{\varepsilon} + \varepsilon \\
&\lesssim \sqrt{\varepsilon},
\end{aligned}$$

where we have used $k^{-1/12} \leq \varepsilon$, and

$$\int \theta^k_{A_0 \cap S} + \theta^k_{A_1 \cap S} \, d\nu_{\text{large}} \leq \int \theta^k_{A_0 \cap S} + \theta^k_{A_1 \cap S} \, d\nu \leq O(\varepsilon),$$

by (5), and the fact that

$$\int d\nu_{\text{large}} = O(1).$$

The latter fact follows because ν_{large} is supported on sets S with $\lambda_2(A_0 \cap S) + \lambda_2(A_1 \cap S) \geq 1/64$. By the isoperimetric inequality in the plane and the fact that F is 1-Lipschitz, we conclude that $\int \nu_{\text{large}} = O(1)$.

We use Theorem 2.5 to bound,

$$\begin{aligned}
\int \rho \, d\nu_{\text{small}} &\lesssim \int \theta^k_{S \cap A_0} + \theta^k_{S \cap A_1} + k^{-1/6} \, d\nu_{\text{small}} \\
&\leq \int \theta^k_{S \cap A_0} + \theta^k_{S \cap A_1} \, d\nu + \varepsilon^2 \int d\nu_{\text{small}} \\
&\lesssim \varepsilon,
\end{aligned}$$

where we have used (5) and the fact that $\int d\nu_{\mathrm{small}} \lesssim \frac{1}{\sqrt{\varepsilon}}$. The latter fact is again by the isoperimetric inequality in $\mathbb{R}^2$, the fact that F is 1-Lipschitz, and the assumption that ν_{small} is supported on sets of measure at least ε.

A final application of the isoperimetric inequality in $\mathbb{R}^2$ shows that since ν_{tiny} is supported on sets of measure at most $O(\varepsilon)$,

$$\int \rho(S)\, d\nu_{\mathrm{tiny}}(S) \leq \int \lambda_2(S \cap A_0) + \lambda_2(S \cap A_1)\, d\nu_{\mathrm{tiny}}(S) \lesssim \varepsilon \cdot \frac{1}{\sqrt{\varepsilon}} = \sqrt{\varepsilon}.$$

It now follows that $D \gtrsim \frac{1}{\sqrt{\varepsilon}} \gtrsim \sqrt{\frac{r}{\log r}}$. On the other hand, we started with a space on $|N| = O(2^r \delta^{-2}) = O(2^{2r} D)$ points, hence we conclude that $D \gtrsim \sqrt{\frac{\log n}{\log \log n}}$, completing the proof.

Extending to the Laakso-fold. The only complication in extending the preceding argument to the Laakso-fold is that, at any given scale, only 1/4 of the dyadic squares correspond to an induced sphere (as opposed to all the squares in the diamondfold). This means that the "monochromatic sheet principle" does not immediately apply. However, the following straightforward strengthening of Theorem 2.2 says that *most dyadic squares are good,* not just one of them. After finding an induced sphere on which both sheets satisfy the theorem, the proof proceeds without change.

Theorem 2.6. *For every metric space Y and 1-Lipschitz mapping $F : [0,1]^2 \to Y$, every $0 < \eta < 1$, and every $0 < \varepsilon < 1$, at least a $(1-\eta)$-fraction of dyadic squares $A \in \mathcal{D}$ of length at least 2^{-r}, with $r \lesssim \frac{1}{\varepsilon} \log \frac{k}{\varepsilon}$, satisfy*

$$\int \mathbb{E}\left[\Delta_F(A[\ell])\right] d\mu(\ell) \leq \frac{\varepsilon}{\eta} 2^{-2r} + \int \Delta_F(\partial A \cap \ell)\, d\mu(\ell). \tag{6}$$

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

[ARV04] Sanjeev Arora, Satish Rao, and Umesh Vazirani. Expander flows, geometric embeddings and graph partitioning. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 222–231 (electronic), New York, 2004. ACM.

[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.

[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.*

[GKL03] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *44th Symposium on Foundations of Computer Science*, pages 534–543, 2003.

[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.

[KR06] Robert Krauthgamer and Yuval Rabani. Improved lower bounds for embeddings into L_1. In *SODA '06: Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 1010–1017, New York, NY, USA, 2006. ACM Press.

[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.

[Laa02] Tomi J. Laakso. Plane with A_∞-weighted metric not bi-Lipschitz embeddable to $\mathbb{R}^N$. *Bull. London Math. Soc.*, 34(6):667–676, 2002.

[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.

[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.

[LM10] J. R. Lee and M. Moharrami. Bilipschitz snowflakes and metrics of negative type. In *42nd Annual ACM Symposium on the Theory of Computing*. ACM, 2010.

[LN04] J. R. Lee and A. Naor. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geom. Funct. Anal.*, 14(4):745–747, 2004.

[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.

[LP01] Urs Lang and Conrad Plaut. Bilipschitz embeddings of metric spaces into space forms. *Geom. Dedicata*, 87(1-3):285–307, 2001.

[LR99] T. Leighton and S. 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.

[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.

[NR03] Ilan Newman and Yuri Rabinovich. A lower bound on the distortion of embedding planar metrics into Euclidean space. *Discrete Comput. Geom.*, 29(1):77–81, 2003.

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

[Pau01] Scott D. Pauls. The large scale geometry of nilpotent Lie groups. *Comm. Anal. Geom.*, 9(5):951–982, 2001.

[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.

A Appendix

A.1 The diamondfold and the Laakso-fold

We first describe our two main constructions. Let X be a separable metric space. An *identified square* in X is an injective mapping $h : [0,1]^2 \to X$ which has distortion 1 (thus, it is an isometry composed with a dilation).

We define a new metric space $Z = \oplus(X, h)$ as follows. Let $Q = h(\partial[0,1]^2)$. Let A be isometric to $h([0,1]^2)$, and disjoint from X, and let $\pi : A \to Q$ be some isometry. Then we define the quotient metric $Z = (X \uplus A)/\sim$, where $\uplus$ represents the disjoint union, and $\sim$ is the equivalence relation given by $x \sim y$ if and only if $x \in A, y \in X$ and $y = \pi(x)$. In words, we glue a new square into X along the boundary $h(\partial[0,1]^2)$.

Now, suppose that S is a finite set of identified squares in X which are interior disjoint, in the sense that for any $h, h' \in S$, we have $h([0,1]^2) \cap h'([0,1]^2) \subseteq h(\partial[0,1]^2) \cap h'(\partial[0,1]^2)$. We write $\oplus(X, S)$ for the space defined as follows. Enumerate $S = \{h_1, h_2, \ldots h_k\}$. Put $X_0 = X$ and for $i = 0, 1, \ldots, k-1$, $X_{i+1} = \oplus(X_i, h_i)$, where we think of h_i as an identified square in X_i in the obvious way. It is easy to see that for a collection of interior disjoint squares S, the final space X_k is, up to isometry, independent of the sequence of gluings. We set $\oplus(X, S) = X_k$. We will use the notation $\mathsf{sq}(X, S)$ to denote the set of $2|S|$ identified squares in $\oplus(X, S)$ given by the original squares in S, along with the $|S|$ newly created squares.

Finally, if h is an identified square in X, we define the set of four identified squares

$$\mathsf{sq}_4(X, h) = \left\{ h|_{[0,\frac{1}{2}]^2}, h|_{[\frac{1}{2},1]^2}, h|_{[0,\frac{1}{2}]\times[\frac{1}{2},1]}, h|_{[\frac{1}{2},1]\times[0,\frac{1}{2}]} \right\}.$$

For a set S of identified squares in X, we use the notation $\mathsf{sq}_4(S) = \bigcup_{h\in S} \mathsf{sq}_4(X, h)$.

The diamondfolds. We will now define a sequence of metric spaces $\{\mathcal{X}_k\}_{k\geq 0}$. We will also keep track of a set of identified squares S_k for each such space. For each $k \geq 0$, $\mathcal{X}_k$ will carry a canonical isometric injection into $\mathcal{X}_{k+1}$, and thus identified squares in $\mathcal{X}_k$ can be thought of as identified squares in $\mathcal{X}_{k+1}$.

Put $\mathcal{X}_0 = [0,1]^2$ and $S_0 = \mathcal{X}_0$. Now, we define $\mathcal{X}_{k+1} = \oplus(\mathcal{X}_k, \mathsf{sq}_4(S_k))$, and set $S_{k+1} = \mathsf{sq}(\mathcal{X}_k, \mathsf{sq}_4(S_k))$.

The Laakso-folds. For the Laakso-fold sequence, we need two additional operations For a space X and an identified square h, define the identified square

$$\mathsf{inn}(X,h) = h|_{[\frac{1}{4},\frac{3}{4}]^2}.$$

For a set S of identified squares in X, we put $\mathsf{inn}(S) = \bigcup_{h\in S} \mathsf{inn}(X,h)$. Furthermore, we define an sq_{16} operator, akin to the sq_4 operator above.

Again, we define a sequence of metric spaces $\{\mathcal{L}_k\}_{k\geq 0}$, along with a collection of identified squares S_k. We put $\mathcal{L}_0 = [0,1]^2$ and $S_0 = \mathcal{L}_0$. Then, $\mathcal{L}_{k+1} = \oplus(\mathcal{L}_k, \mathsf{inn}(S_k))$, and $S_{k+1} = \mathsf{sq}_{16}(S_k) \cup \mathsf{sq}(\mathcal{L}_k, \mathsf{inn}(S_k))$.

B A differentiation argument

We begin with some general definitions. Let X and Y be arbitrary metric spaces. For any function $F : X \to Y$ and any finite sequence $s = \langle s_1, s_2, \ldots, s_j\rangle$ of points in X, define

$$\Delta_F(s) = \sum_{i=1}^{j-1} d_Y(F(s_i), F(s_{i+1})).$$

Let $s = \langle s_1, s_2, \ldots, s_j\rangle$ and $t = \langle t_1, t_2, \ldots, t_k\rangle$ be two sequences of points in X. We will use $|s|$ to denote the number of terms in such a sequence. Say that a map $g : [j] \to [k]$ is *increasing* if $g(i) \leq g(i+1)$ for all $i = 1, 2, \ldots, j-1$. We define an (asymmetric) notion of closeness between sequences by,

$$\rho(s;t) = \inf_{\substack{g:[j]\to[k] \\ \text{increasing}}} \sup_{i\in[j]} d(s_i, t_{g(i)}).$$

Lemma B.1. *If s and t are two sequences in X, then*

$$\Delta_F(s) \leq \Delta_F(t) + 2|s| \cdot \|F\|_{\mathrm{Lip}} \cdot \rho(s;t).$$

Proof. Let $g : [|s|] \to [|t|]$ achieve the distance $\rho(s;t)$. Then,

$$\begin{aligned}
\Delta_F(s) &= \sum_{i=1}^{|s|-1} d_Y(F(s_i), F(s_{i+1})) \\
&\leq \sum_{i=1}^{|s|-1} \left[d_Y\left(F\left(t_{g(i)}\right), F\left(t_{g(i+1)}\right)\right) + d_Y\left(F\left(s_i\right), F\left(t_{g(i)}\right)\right) + d_Y\left(F\left(s_{i+1}\right), F\left(t_{g(i+1)}\right)\right)\right] \\
&\leq \Delta_F(t) + \|F\|_{\mathrm{Lip}} \sum_{i=1}^{|s|-1} d_Y\left(s_i, t_{g(i)}\right) + d_Y\left(s_{i+1}, t_{g(i+1)}\right) \\
&\leq \Delta_F(t) + \|F\|_{\mathrm{Lip}} \cdot 2|s|\,\rho(s;t).
\end{aligned}$$

$\square$

B.1 Measuring variation on grates

We first specialize to the situation $X = [0,1]$. We will call P a *subdivision* of $[0,1]$ if P is a cover of $[0,1]$ by closed intervals which intersect only at their endpoints. For a subdivision P of $[0,1]$, we write $\partial P = \bigcup_{s\in P} \partial s$. A *grate on* $[0,1]$ is a pair $g = (g_P, g_S)$ where P is a finite subdivision of $[0,1]$, and $S \subseteq [0,1]$ is a finite subset with $\partial P \subseteq S$. Given a grate g, we use the notation $\mathcal{P}(g)$ for g_P and $\mathcal{S}(g)$ for g_S. We call a sequence of grates $g_0, g_1, \ldots, g_k$ a *nested grating* if $\mathcal{P}(g_i)$ is a refinement of $\mathcal{P}(g_{i-1})$ for every $i = 1, 2, \ldots, k$. We also think of any finite subset $s \subseteq [0,1]$ as a sequence endowed with the induced ordering.

Lemma B.2. *For any metric space Y, any mapping $F : [0,1] \to Y$, and any finite nested grating $g_0, g_1, \ldots, g_k$, we have*

$$\begin{aligned}\sum_{i=1}^{k} \sum_{s\in\mathcal{P}(g_i)} &\left[\Delta_F(s\cap\mathcal{S}(g_i)) - \Delta_F(\partial s)\right] \\ &= \Delta_F(\mathcal{S}(g_k)) - \Delta_F(\mathcal{S}(g_0)) + \sum_{i=0}^{k-1} \sum_{s\in\mathcal{P}(g_i)} \left[\Delta_F(s\cap\mathcal{S}(g_i)) - \Delta_F(s\cap\partial P(g_{i+1}))\right]. \end{aligned} \tag{7}$$

Proof. First, observe that

$$\sum_{i=1}^{k} \sum_{s\in\mathcal{P}(g_i)} \Delta_F(s\cap\mathcal{S}(g_i)) = \sum_{i=1}^{k} \Delta_F(\mathcal{S}(g_i)) = \Delta_F(\mathcal{S}(g_k)) - \Delta_F(\mathcal{S}(g_0)) + \sum_{i=0}^{k-1} \sum_{s\in\mathcal{P}(g_i)} \Delta_F(s\cap\mathcal{S}(g_i)).$$

Subtracting the left and right-hand side from the respective sides of (7), we are left to prove that,

$$\sum_{i=1}^{k} \sum_{s\in\mathcal{P}(g_i)} \Delta_F(\partial s) = \sum_{i=0}^{k-1} \sum_{s\in\mathcal{P}(g_i)} \Delta_F(s\cap\partial P(g_{i+1})),$$

which follows from the identity, for every $i = 1, 2, \ldots, k$,

$$\sum_{s\in\mathcal{P}(g_i)} \Delta_F(\partial s) = \sum_{s\in P(g_{i-1})} \Delta_F(s\cap\partial P(g_i)),$$

because $\mathcal{P}(g_i)$ is a refinement of $\mathcal{P}(g_{i-1})$. □

Lemma B.3. *Let $g_0, g_1, \ldots, g_k$ be a finite nested grating such that for each $i = 0, 1, \ldots, k$, $\partial P(g_i)$ is κ_i-dense in $[0,1]$. Then for any metric space Y, and any Lipschiz mapping $F : [0,1] \to Y$, we have*

$$\sum_{i=1}^{k} \sum_{s\in\mathcal{P}(g_i)} \left[\Delta_F(s\cap\mathcal{S}(g_i)) - \Delta_F(\partial s)\right] \le \|F\|_{\mathrm{Lip}} \left(1 + 4\sum_{i=0}^{k-1} |\mathcal{S}(g_i)|\,\kappa_{i+1}\right).$$

Proof. First, we have $\Delta_F(\mathcal{S}(g_k)) \le \|F\|_{\mathrm{Lip}}$, by definition. Hence by Lemma B.2,

$$\sum_{i=1}^{k} \sum_{s\in\mathcal{P}(g_i)} \left[\Delta_F(s\cap\mathcal{S}(g_i)) - \Delta_F(\partial s)\right] \le \|F\|_{\mathrm{Lip}} + \sum_{i=0}^{k-1} \sum_{s\in\mathcal{P}(g_i)} \left[\Delta_F(s\cap\mathcal{S}(g_i)) - \Delta_F(s\cap\partial P(g_{i+1}))\right]. \tag{8}$$

Next, we have the estimate, for any $i = 0, 1, \dots, k-1$ and $s \in \mathcal{P}(g_i)$,

$$\rho\left(s \cap \mathcal{S}(g_i)\right); s \cap \partial P(g_{i+1})) \leq \kappa_{i+1},$$

since $\partial P(g_{i+1})$ is κ_{i+1}-dense in $[0,1]$. Appealing to Lemma B.1, it follows that,

$$\Delta_F(s \cap \mathcal{S}(g_i)) - \Delta_F(s \cap \partial P(g_{i+1})) \leq 2|s \cap \mathcal{S}(g_i)| \cdot \|F\|_{\mathrm{Lip}} \cdot \kappa_{i+1},$$

and thus,

$$\sum_{s \in \mathcal{P}(g_i)} \Delta_F(s \cap \mathcal{S}(g_i)) - \Delta_F(s \cap \partial P(g_{i+1})) \leq 4|\mathcal{S}(g_i)| \cdot \|F\|_{\mathrm{Lip}} \cdot \kappa_{i+1}.$$

Summing over $i = 0, 1, \dots, k-1$ and applying (8) yields the statement of the lemma. $\square$

B.2 Average variation along lines

We now consider the unit square $[0,1]^2$. For every line $\ell \subseteq \mathbb{R}^2$, fix an arbitrary orientation so that a subset $s \subseteq \ell$ can naturally be thought of as a sequence. We will call $A \subseteq [0,1]^2$ a *dyadic square of side length* 2^{-r} if it is a closed square whose corners are integer multiples of 2^{-r}. We will use $\mathcal{D}_r$ to denote the set of all dyadic squares of side length 2^{-r}, and write $\mathcal{D} = \bigcup_{r=0}^{\infty} \mathcal{D}_r$. We use $A_0 = [0,1]^2$ to denote the dyadic square of side length 1.

For every dyadic square A, let $A[\ell]$ be an arbitrary finite subset $A[\ell] \subseteq A \cap \ell$ which contains $\partial A \cap \ell$. For a dyadic square $A \in \mathcal{D}_r$ and a number $h \in \mathbb{N}$, we write

$$\mathsf{sub}_h(A) = \{B \subseteq A : B \in \mathcal{D}_{r+h}\},$$

and we define

$$\partial\, \mathsf{sub}_h(A) = \bigcup_{B \in \mathsf{sub}_h(A)} \partial B.$$

We observe that the definition of a grating extends to any interval, without change. The above structure yields, for every line ℓ, a nested grating $g_0, g_1, \dots$ of $\ell \cap [0,1]^2$ as follows. Let $h \in \mathbb{N}$ be a parameter. We construct the *h-grating* as follows: $\mathcal{P}(g_0) = A_0 \cap \ell$ and $\mathcal{S}(g_0) = A_0[\ell]$. In general, $\mathcal{P}(g_i) = \bigcup_{A \in \mathcal{D}_{ih}} (A \cap \ell)$ and $\mathcal{S}(g_i) = \bigcup_{A \in \mathcal{D}_{ih}} A[\ell]$. The next proposition is straightforward.

Proposition B.4. *If $g_0, g_1, \dots$ is the h-grating of $\ell \cap [0,1]^2$ specified above, then $\partial \mathcal{P}(g_i)$ is $\sqrt{2} \cdot 2^{-hi}$-dense in $\ell \cap [0,1]^2$, and*

$$|\mathcal{S}(g_i)| \leq 2 \cdot 2^{ih} \max_{A \in \mathcal{D}_{ih}} |A[\ell]|.$$

The next corollary follows from Lemma B.3 and Proposition B.4.

Corollary B.5. *For any line $\ell \subseteq \mathbb{R}^2$, any $m \in \mathbb{N}$, and any h-grating $g_0, g_1, \dots$ of $\ell \cap [0,1]^2$, the following holds. If $k = \max_{A \in \mathcal{D}} A[\ell]$, Y is a metric space and $F : [0,1]^2 \to Y$ is a 1-Lipschitz mapping, then*

$$\sum_{i=1}^{m} \sum_{A \in \mathcal{D}_{hi}} \Delta_F(A[\ell]) - \Delta_F(\partial A \cap \ell) \leq \sqrt{2} + 8\sqrt{2} k m 2^{-h}.$$

Finally, by taking a family of gratings, one for each line $\ell \subseteq \mathbb{R}^2$, we obtain the following.

Theorem B.6. *For any $0 < \varepsilon < 1$, and any metric space Y, the following holds. Suppose that for every $A \in \mathcal{D}$ and line $\ell \subseteq \mathbb{R}^2$, we have a finite set $A[\ell] \subseteq A \cap \ell$ which contains $\partial A \cap \ell$. If $k = \sup_{A\in\mathcal{D},\ell} |A[\ell]| < \infty$, and $F : [0,1]^2 \to Y$ is any 1-Lipschitz mapping, then there is a dyadic square A of side length 2^{-r}, with $r \lesssim \frac{1}{\varepsilon}\log\frac{k}{\varepsilon}$, and such that*

$$\int \Delta_F(A[\ell])\, d\mu(\ell) \le \varepsilon 2^{-2r} + \int \Delta_F(\partial A \cap \ell)\, d\mu(\ell). \tag{9}$$

Proof. From Corollary B.5, we have, for every $h, m \in \mathbb{N}$,

$$\begin{aligned}\sum_{i=1}^{m} \sum_{A\in\mathcal{D}_{hi}} \int \Delta_F(A[\ell]) - \Delta_F(\partial A \cap \ell)\, d\mu(\ell) &= \int \sum_{i=1}^{m} \sum_{A\in\mathcal{D}_{hi}} \Delta_F(A[\ell]) - \Delta_F(\partial A \cap \ell)\, d\mu(\ell) \\ &\le O(1 + km2^{-h}).\end{aligned}$$

Thus, there exists an $i \le m$ for which,

$$\sum_{A\in\mathcal{D}_{hi}} \int \Delta_F(A[\ell]) - \Delta_F(\partial A \cap \ell)\, d\mu(\ell) \le O(k2^{-h} + 1/m).$$

Choosing some $m = O(1/\varepsilon)$ and $h = O(\log(k/\varepsilon))$ yields an $A \in D_{hi}$ for which

$$\int \Delta_F(A[\ell]) - \Delta_F(\partial A \cap \ell)\, d\mu(\ell) \le \varepsilon 2^{-2hi}.$$

Thus (9) holds for this square A, and clearly the side length of A is at least 2^{-2hi}, where $hi = O(hm) = O(\frac{1}{\varepsilon}\log\frac{k}{\varepsilon})$, completing the proof. ☐

In the coming sections, it will be necessary to choose the points of $A[\ell]$ randomly, thus we record the following corollary. For intuition, the reader may wish to think about the distribution we use in the sections to come: For some parameter k, $A[\ell]$ is chosen to be $S \cup (\partial A \cap \ell)$, where $S \subseteq A \cap \ell$ is a uniformly random subset of size k.

Corollary B.7. *Suppose that for every $A \in \mathcal{D}$ and every line $\ell \subseteq \mathbb{R}^2$, we have a* random finite set *$A[\ell]$ of size at most k. Then for every metric space Y and 1-Lipschitz mapping $F : [0,1]^2 \to Y$, and every $0 < \varepsilon < 1$, there exists a dyadic square A of side length at least 2^{-r}, with $r \lesssim \frac{1}{\varepsilon}\log\frac{k}{\varepsilon}$, and such that*

$$\int \mathbb{E}\left[\Delta_F(A[\ell])\right] d\mu(\ell) \le \varepsilon 2^{-2r} + \int \Delta_F(\partial A \cap \ell)\, d\mu(\ell). \tag{10}$$

C The symmetric difference of parallel slices

Recall that Λ is the set of all lines in $\mathbb{R}^2$ which intersect $[0,1]^2$. For a line $\ell \in \Lambda$ we define the *angle* of ℓ to be the angle formed by ℓ and the x-axis. For any $\theta \in [0, 2\pi)$, let Λ_θ be the set of lines in Λ of angle θ. Let μ_θ be the restriction of μ on Λ_θ.

Let $\ell \subset \mathbb{R}^2$ be a line intersecting $[0,1]^2$. Let $k > 0$. We define a distribution $\mathcal{P}(\ell, k)$ of sequences of points $(p_0, \ldots, p_{k+1})$, where for every $i \in \{0, \ldots, k+1\}$, $p_i \in \mathbb{R}^2$. We first set p_0 and p_k to be the two end-points of the segment $\ell \cap [0,1]^2$. We choose a set Q of k points in s, uniformly and independently at random, and for every $i \in \{1, \ldots, k\}$, we set p_i to be the i-th point in Q closest to p_0.

If $\ell \subset \mathbb{R}^2$ is a line that does not intersect $[0,1]^2$, then we set $\mathcal{P}(\ell, k)$ to be the trivial distribution which always returns $(q, \ldots, q) \in (\mathbb{R}^2)^{k+2}$, for some arbitrary point $q \notin [0,1]^2$.

Definition C.1 (Complexity of a set)**.** *Let $S \subseteq [0,1]^2$, $\eta > 0$. We define the η-*complexity *of S, denoted by $\mathcal{C}_\eta(S)$ to be*

$$\mathcal{C}_\eta(S) = \int_\Lambda \mathbb{E}_{P \in \mathcal{P}(\ell, 1/\eta)} \sum_{j=0}^{1/\eta} \left| \mathbf{1}_S(p_j) - \mathbf{1}_S(p_{j+1}) \right| d\mu(\ell).$$

We also define the boundary-complexity *of S, denoted by $\mathcal{C}(S)$ to be*

$$\mathcal{C}(S) = \int_\Lambda \mathbb{E}_{P \in \mathcal{P}(\ell, 0)} \left| \mathbf{1}_S(p_0) - \mathbf{1}_S(p_1) \right| d\mu(\ell).$$

C.1 The structure of low complexity sets

For the remainder of this section, let $\varepsilon \in \left(0, \frac{1}{64}\right)$, $\eta = \frac{1}{2} \cdot \varepsilon^6$.

Definition C.2 (ε-Low complexity set)**.** *We say that a set $S \subseteq [0,1]^2$ is of ε-*low complexity *if*

$$|\mathcal{C}_\eta(S) - \mathcal{C}(S)| \leq 2^{-27} \cdot \varepsilon^2.$$

The following Lemma is immediate. We include the proof for completeness.

Lemma C.3. *Let $A \subseteq [0,1]^2$ be a polygonal set, and let $\ell \subset \mathbb{R}^2$ be a line such that $\lambda_1(\ell \cap A) \geq \varepsilon^6$. Then, $\mathbb{P}_{P \in \mathcal{P}(\ell, 1/\eta)}[\exists i \in \{0, \ldots, 1/\eta + 1\}$ s.t. $p_i \in A] > 1/2$.*

Proof. We have

$$\begin{aligned}
\mathbb{P}_{P \in \mathcal{P}(\ell, 1/\eta)}[\exists i \in \{0, \ldots, 1/\eta + 1\} \text{ s.t. } p_i \in A] &\geq \mathbb{P}_{P \in \mathcal{P}(\ell, 1/\eta)}[\exists i \in \{1, \ldots, 1/\eta\} \text{ s.t. } p_i \in A] \\
&= 1 - \mathbb{P}_{P \in \mathcal{P}(\ell, 1/\eta)}[\forall i \in \{1, \ldots, 1/\eta\}, p_i \notin A] \\
&= 1 - \left(1 - \frac{\lambda_1(A \cap \ell)}{\lambda_1([0,1]^2 \cap \ell)}\right)^{1/\eta} \\
&\geq 1 - \left(1 - \frac{\varepsilon^6}{\sqrt{2}}\right)^{1/\eta} \\
&= 1 - \left[\left(1 - \frac{\varepsilon^6}{\sqrt{2}}\right)^{\frac{\sqrt{2}}{\varepsilon^6}}\right]^{\sqrt{2}} \\
&> 1/2.
\end{aligned}$$

$\square$

We will use the following elementary geometric fact.

Lemma C.4. *Let $\beta > 0$. Let $x_1, x_2 \in \partial[0,1]^2$, and let $A_3, A_4 \subset \partial[0,1]^2$ be line segments. For any $i \in \{1, 2\}$, let $X_i = \bigcup_{y \in A_{i+2}} x_i y$. Suppose that the following conditions are satisfied:*

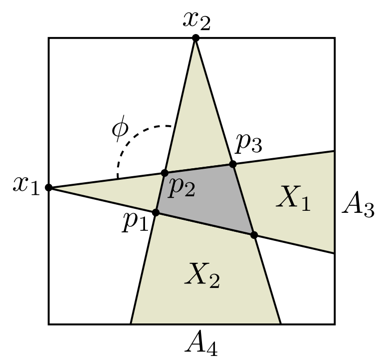
(i) The sets $\{x_1\}, \{x_2\}, A_3, A_4$ are pairwise disjoint, and they appear in the order they are listed along a clockwise traversal of $\partial[0,1]^2$.

(ii) For any $i \in \{1, 2\}$, the point x_i and the set A_{i+2} are in different sides of $\partial[0,1]^2$.

(iii) For any $y_1 \in X_1$, we have $\|y - x_2\|_2 \geq \beta$, and for any $y_2 \in X_2$, we have $\|y - x_1\|_2 \geq \beta$.

Then, $\lambda_2(X_1 \cap X_2) \geq \frac{\lambda_1(A_3)\cdot\lambda_1(A_4)\cdot\beta^5}{64\cdot\sqrt{2}}$.

Proof. Consider the points p_1, p_2, p_3 as in the following figure.



Note that in general, the above figure can look slightly different. For example, the points x_1 and x_2 can be on the same side of $\partial[0,1]^2$. However, the relative position of all points remains topologically the same, so we can fix w.l.o.g. the notation for p_1, p_2, p_3.

Let $\phi \in [0, \pi]$ be the angle formed by $x_2 p_2 x_1$. By condition (iii) it follows that

$$\sin\phi > \frac{\beta}{2\sqrt{2}}.$$

Similarly, condition (iii) implies

$$\|p_1 - p_2\|_2 \geq \frac{\lambda_1(A_3)\cdot\beta^2}{4}, \text{ and } \|p_2 - p_3\|_2 \geq \frac{\lambda_1(A_4)\cdot\beta^2}{4}.$$

Let T be the triangle $p_1 p_2 p_3$. It follows that

$$\lambda_2(X_1 \cap X_2) > \lambda_2(T) = \frac{1}{2}\cdot\|p_1 - p_2\|_2\cdot\|p_2 - p_3\|_2\cdot\sin\phi > \frac{\lambda_1(A_3)\cdot\lambda_1(A_4)\cdot\beta^5}{64\cdot\sqrt{2}},$$

as required. □

We next derive an obstruction to ε-low complexity.

Lemma C.5 (Conflicting segments)**.** *Let $S \subseteq [0,1]^2$, and let $I_1, I_2, I_3, I_4 \subset \partial[0,1]^2$ such that the following conditions are satisfied:*

- *For any $i \in \{1, \ldots, 4\}$, the set I_i is a finite collection of disjoint line segments.*
- *For any $i \neq j \in \{1, \ldots, 4\}$, we have $I_i \cap I_j = \emptyset$.*
- *For any $i \in \{1, \ldots, 4\}$, we have $\lambda_1(I_i) \geq 2^{-11}\cdot\varepsilon$.*

- $I_1 \cup I_3 \subseteq S$.
- $I_2 \cup I_4 \subseteq [0,1]^2 \setminus S$.
- *The sets* I_1, I_2, I_3, I_4 *appear in the order they are listed, along a clock-wise traversal of* $\partial[0,1]^2$. *Formally, there exists a homeomorphism* $\phi : \partial[0,1]^2 \to \{(\cos x, \sin x) : x \in [0, 2\pi)\}$, *where* S_1 *denotes the unit circle, such that for any* $p_1 \in I_1, \ldots, p_4 \in I_4$, *if* $\phi(p_i) = (\cos x_i, \sin x_i)$, $x_i \in [0, 2\pi)$, *for any* $i \in \{1, \ldots, 4\}$, *then* $x_1 < x_2 < x_3 < x_4$.
- *For any* $i \in \{1, 2\}$, *the sets* I_i *and* I_{i+2} *are contained in different sides of* $\partial[0,1]^2$.

Then, S *is not of* ε*-low complexity.*

Proof. For any $x_1 \in I_1$, $x_3 \in I_3$, we say that the segment x_1x_3 is *good*, if $\lambda_1(x_1x_3 \setminus S) < \varepsilon^6$, and otherwise we say that it is *bad* (see Figure 6 for an example). We say that $x_1 \in I_1$ is *illuminating* if

$$\lambda_1(\{x_3 \in I_3 : x_1x_3 \text{ is good}\}) \geq \lambda_1(I_3)/2.$$

Notice that by Lemma C.3, we have

$$|\mathcal{C}_\eta(S) - \mathcal{C}(S)| > \lambda_1(\{x_1 \in I_1 : x_1 \text{ is not illuminating}\}) \cdot \lambda_1(I_3)/8.$$

Since S is of ε-low complexity, it follows that

$$\lambda_1(\{x_1 \in I_1 : x_1 \text{ is not illuminating}\}) < \frac{8 \cdot 2^{-27}}{2^{-22}} \lambda_1(I_1) < \frac{1}{2} \lambda_1(I_1).$$

Thus,

$$\lambda_1(\{x_1 \in I_1 : x_1 \text{ is illuminating}\}) > \frac{1}{2} \lambda_1(I_1).$$

Therefore, there exists an illuminating $x_1^* \in I_1$, such that the distance between x_1^* and the first and last points in I_1 along a clockwise traversal of $\partial[0,1]^2$, is at least $\lambda_1(I_1)/4$.

Let

$$J_3 = \{x_3 \in I_3 : x_1^* x_3 \text{ is good}\}.$$

Since x_1^* is illuminating, we have

$$\lambda_1(J_3) \geq \lambda_1(I_3)/2 \geq 2^{-12}\varepsilon.$$

Let I_3' be the central part of J_3 of measure

$$\lambda_1(I_3') = \lambda_1(J_3)/2 \geq 2^{-13}\varepsilon.$$

That is, the total measure of points in J_3 preceeding I_3' is exactly $\lambda_1(J_3)/4$. Note that for any $x_2 \in I_2$, and $x_3 \in I_3'$, the distance between the segment $x_1^* x_3$ and the x_2, is at least $2^{-27} \cdot \varepsilon^2$.

For any $i \in \{2, 4\}$, let I_i' be the central part of I_i, of total measure $\lambda_1(I_i') = \lambda_1(I_i)/2$. It follows that for any $x_2 \in I_2'$, $x_4 \in I_4'$, the distance between x_1^* and the segment x_2x_4 is at least $2^{-26} \cdot \varepsilon^2$.

Consider now some $x_2 \in I_2'$. Let

$$J_4(x_2) = \{x_4 \in I_4' : \lambda_1(x_2x_4 \cap S) \geq \varepsilon^6\}.$$

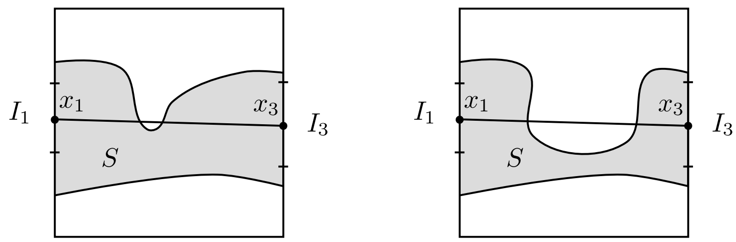


Figure 6: Example of good (left) and bad (right) segments x_1x_3.

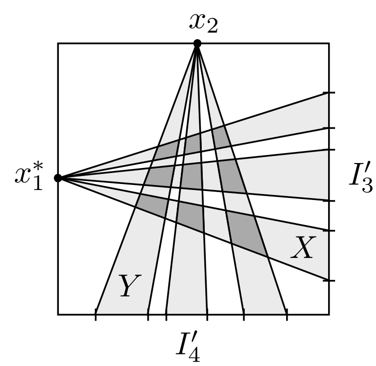


Figure 7: The sets X and Y in the proof of Lemma C.5.

Let $X = \bigcup_{x_3 \in I_3'} x_1^* x_3$, and $Y = \bigcup_{x_4 \in I_4'} x_2 x_4$. Each of the sets X, Y is the union of finitely many triangles $X = X_1 \cup \ldots \cup X_t$, $Y = Y_1 \cup \ldots \cup Y_s$ (see Figure 7). For each $i \in \{1, \ldots, t\}$, let A_i^X be the side of the triangle X_i that does not contain x_1^*. Similarly, for each $j \in \{1, \ldots, s\}$, let A_j^Y be the side of the triangle Y_j that does not contain x_2. Observe that $I_3' = \bigcup_{i=1}^t A_i^X$, and $I_4' = \bigcup_{j=1}^s A_j^Y$. Using Lemma C.4, we deduce that for every $i \in \{1, \ldots, t\}$, and for any $j \in \{1, \ldots, s\}$, we have

$$\lambda_2(X_i \cap Y_j) \geq \frac{\lambda_1(A_i^X) \cdot \lambda_1(A_j^Y)}{2^{16} \cdot \sqrt{2}} \cdot 2^{-44} \cdot \varepsilon^4.$$

Thus,

$$\begin{aligned} \lambda_2(X \cap Y) &= \lambda_2 \left(\bigcup_{i=1}^{t} \bigcup_{j=1}^{s} (X_i \cap Y_j) \right) \\ &= \sum_{i=1}^{t} \sum_{j=1}^{s} \lambda_2(X_i \cap Y_j) \\ &\geq \frac{2^{-44} \cdot \varepsilon^4}{2^{16} \cdot \sqrt{2}} \cdot \lambda_1(I_3') \cdot \lambda_1(I_4') \\ &\geq \frac{2^{-48} \cdot \varepsilon^6}{\sqrt{2}}. \end{aligned} \tag{11}$$

Since for every $x_3 \in I'_3$, the segment $x_1^* x_3$ is good, it follows that

$$\lambda_2(X \setminus S) < \varepsilon^6 \cdot \lambda_1(I'_3) < \varepsilon^6. \tag{12}$$

From (12) we get

$$\lambda_2((X \cap Y) \setminus S) \leq \lambda_2(X \setminus S) < \varepsilon^6. \tag{13}$$

By (11) and (13), we get

$$\lambda_2(X \cap Y \cap S) > \lambda_2(X \setminus S) \cdot \frac{15}{16}. \tag{14}$$

By (14) it easily follows that for any $x_2 \in I'_2$, we have $\lambda_1(J_4(x_2)) \geq \lambda_1(I'_4)/2$. By Lemma C.3, this implies that

$$|\mathcal{C}_\eta(S) - \mathcal{C}(S)| \geq \frac{1}{4} \cdot \lambda_1(I'_2) \cdot \lambda_1(I'_4) > 2^{-27} \cdot \varepsilon^2,$$

contradicting the fact that S is of ε-low complexity, and concluding the proof. $\square$

Let $\Gamma_0, \ldots, \Gamma_3$ be the four sides of the square $\partial[0,1]^2$, and suppose they appear in the order they are listed along a clock-wise traversal of $\partial[0,1]^2$, and we write Γ_i for $\Gamma_{i \bmod 4}$.

Definition C.6 (Type (I) sets)**.** *Let $S \subseteq [0,1]^2$ be a polygonal set. We say that S is of* type (I) *if there exists $i \in \{0, \ldots, 3\}$, such that if we set $\Gamma' = \Gamma_i \cup \Gamma_{i+1} \cup \Gamma_{i+2}$, then*

$$\lambda_1(\Gamma' \cap S) \geq \lambda_1(\Gamma') - 2^{-7} \cdot \varepsilon,$$

or

$$\lambda_1(\Gamma' \setminus S) \geq \lambda_1(\Gamma') - 2^{-7} \cdot \varepsilon.$$

In the former case we also say that S is of type I-a*, and in the later case we say that it is of type* type I-b*.*

Definition C.7 (Type (II) sets)**.** *Let $S \subseteq [0,1]^2$ be a polygonal set. We say that S is of* type (II) *if there exists a halfplane $H \subset \mathbb{R}^2$, such that*

$$\lambda_1(\partial[0,1]^2 \cap H \cap S) \geq \lambda_1(\partial[0,1]^2 \cap H) - 2^{-7} \cdot \varepsilon,$$

and

$$\lambda_1((\partial[0,1]^2 \setminus H) \setminus S) \geq \lambda_1(\partial[0,1]^2 \setminus H) - 2^{-7} \cdot \varepsilon.$$

Figure 8 depicts examples of sets of type (I) and (II). We remark that a set $S \subseteq [0,1]^2$ can be both of type (I), and type (II), e.g. the set $S = [0,1]^2$.

Lemma C.8 (Classification of ε-low complexity sets)**.** *Let $S \subseteq [0,1]^2$, be a polygonal set of ε-low complexity. Then, S is of type (I), or S is of type (II) (or both).*

Proof. We begin by defining a partition $\mathcal{Z}^1 = \{Z_i^1\}_{i=1}^{|\mathcal{Z}^1|}$ of $\partial[0,1]^2$, into a sequence of consecutive intervals. Let $\phi : [0,4) \to \partial[0,1]^2$ be the mapping that sends every $x \in [0,4)$ to the point $p \in \partial[0,1]^2$ such that the distance between $(0,0)$ and p along a clock-wise traversal of $[0,1]^2$, is x. For any

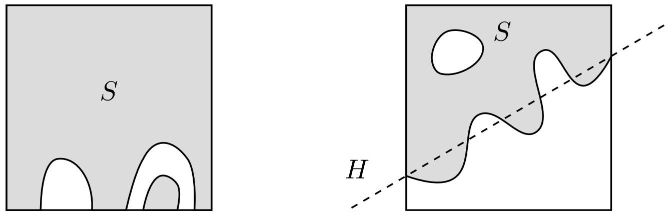


Figure 8: Example of sets of type (I) (left), and type (II) (right).

$0 \leq x < y < 4$, let $I[x,y] = \{\phi(z) : x \leq z < y\}$. We define a sequence $z_0 < \ldots < z_t \in [0,4]$, as follows. Let

$$z_0 = \phi(0) = (0,0).$$

Given z_i, $i \geq 0$, we set

$$z_{i+1} = \sup\{z \in [0,4) : \lambda_1(I[z_0,z] \cap S) \leq 2^{-10} \cdot \varepsilon \text{ or } \lambda_1(I[z_0,z] \setminus S) \leq 2^{-10} \cdot \varepsilon\}.$$

We terminate the sequence at $t = \min\{i : z_i = 4\}$. We set $k_1 = t$, and for any $i \in \{1, \ldots, k_1\}$, we set $Z_i^1 = I[z_{i-1}, z_i]$.

Next we modify $\mathcal{Z}^1$ to obtain a partition $\mathcal{Z}^2 = \{Z_i^2\}_{i=1}^{|\mathcal{Z}^2|}$ of $\partial[0,1]^2$. If for every $Z_i^1 \in \mathcal{Z}^1$, we have $|Z_i^1| \geq 2 \cdot 2^{-10} \cdot \varepsilon$, then we set $\mathcal{Z}^2 = \mathcal{Z}^1$. Otherwise, observe that the only set in $\mathcal{Z}^2$ with total length less than $2 \cdot 2^{-10} \cdot \varepsilon$, is $Z_{|\mathcal{Z}^1|}$. In this case we merge $Z_{|\mathcal{Z}^1|}$ with $Z_{|\mathcal{Z}^1|-1}$. I.e., we set $|\mathcal{Z}^2| = |\mathcal{Z}^1| - 1$, and for any $i < |\mathcal{Z}^2|$, we set $Z_i^2 = Z_i^1$, and $Z^2_{|\mathcal{Z}^2|} = Z^1_{|\mathcal{Z}^1|-1} \cup Z^1_{|\mathcal{Z}^1|}$. This completes the description of $\mathcal{Z}^2$.

Observe that for any $Z_i^2 \in \mathcal{Z}^2$, we have either

$$\lambda_1(Z_i^2 \cap S) \geq 2^{-10} \cdot \varepsilon \quad \text{and} \quad \lambda_1(Z_i^2 \setminus S) \leq 2 \cdot 2^{-10} \cdot \varepsilon,$$

or

$$\lambda_1(Z_i^2 \setminus S) \geq 2^{-10} \cdot \varepsilon \quad \text{and} \quad \lambda_1(Z_i^2 \cap S) \leq 2 \cdot 2^{-10} \cdot \varepsilon,$$

In the former case we say that Z_i^2 is *green*, and in the later case we say that is it *red*. Note that a set Z_i^2 can be both green and red. Observe further that the sequence $\mathcal{Z}^2$ alternates between green and red sets.

We now modify $\mathcal{Z}^2$ to obtain a partition $\mathcal{Z}^3 = \{Z_i^3\}_{i=1}^{|\mathcal{Z}^3|}$ of $\partial[0,1]^2$. If $\mathcal{Z}^2$ contains either a single set, or an even number of sets, then we set $\mathcal{Z}^3 = \mathcal{Z}^2$. Otherwise, we obtain $\mathcal{Z}^3$ from $\mathcal{Z}^2$ by merging the first and last sets of $\mathcal{Z}^2$, and without modifying the remaining sets. We have that for any $Z_i^3 \in \mathcal{Z}^3$, we have either

$$\lambda_1(Z_i^3 \cap S) \geq 2^{-10} \cdot \varepsilon \quad \text{and} \quad \lambda_1(Z_i^3 \setminus S) \leq 4 \cdot 2^{-10} \cdot \varepsilon,$$

or

$$\lambda_1(Z_i^3 \setminus S) \geq 2^{-10} \cdot \varepsilon \quad \text{and} \quad \lambda_1(Z_i^3 \cap S) \leq 4 \cdot 2^{-10} \cdot \varepsilon,$$

As before, in the former case we say that Z_i^3 is *green*, and in the later case we say that is it *red*.

Next, we modify $\mathcal{Z}^3$ to obtain a partition $\mathcal{Z}^4 = \{Z_i^4\}_{i=1}^{|\mathcal{Z}^4|}$ of $\partial[0,1]^2$. For $i = 1, \ldots, |\mathcal{Z}^3|$, we consider the set Z_i^3. If Z_i^3 is contained in a single side Γ_j of the square $\partial[0,1]^2$, then we simply add

Z_i^3 to $\mathcal{Z}^4$. Otherwise, let Γ_j, Γ_k be the first and last sides that intersect Z_i^3, and let $W = \Gamma_j \cap Z_i^3$, $W' = \Gamma_k \cap Z_i^3$. Consider first the case where Z_i^3 is green. If $\lambda_1(W \cap S) \leq 2^{-10} \cdot \varepsilon/4$, then we remove W from Z_i^3, and add it to the current copy of Z_{i-1}^3 in $\mathcal{Z}^4$. Similarly, if $\lambda_1(W' \cap S) \leq 2^{-10} \cdot \varepsilon/4$, then we remove W' from Z_i^3, and add it to the current copy of Z_{i+1}^3, where we use the notation $Z_{|\mathcal{Z}^3|+1}^3 = Z_1^3$. The case where Z_i^3 is red is treated analogously: If $\lambda_1(W \setminus S) \leq 2^{-10} \cdot \varepsilon/4$, then we remove W from Z_i^3, and add it to the current copy of Z_{i-1}^3 in $\mathcal{Z}^4$. Finally, if $\lambda_1(W' \setminus S) \leq 2^{-10} \cdot \varepsilon/4$, then we remove W' from Z_i^3, and add it to the current copy of Z_{i+1}^3. This completes the description of $\mathcal{Z}^4$.

For any $Z_i^4 \in \mathcal{Z}^4$, we have either

$$\lambda_1(Z_i^3 \cap S) \geq \frac{1}{2} 2^{-10} \cdot \varepsilon \quad \text{and} \quad \lambda_1(Z_i^4 \setminus S) \leq \frac{9}{2} \cdot 2^{-10} \cdot \varepsilon,$$

or

$$\lambda_1(Z_i^3 \setminus S) \geq \frac{1}{2} 2^{-10} \cdot \varepsilon \quad \text{and} \quad \lambda_1(Z_i^4 \cap S) \leq \frac{9}{2} \cdot 2^{-10} \cdot \varepsilon,$$

Again, in the former case we say that Z_i^3 is *green*, and in the later case we say that is it *red*. Observe that $\mathcal{Z}^4$ contains either a single set, or an even number of sets, and alternates between green and red sets. Moreover, if a green set $Z \in \mathcal{Z}^4$ intersects a side Γ_j of $\partial[0,1]^2$, then $\lambda_1(Z \cap \Gamma_j \cap S) \geq \frac{1}{4} 2^{-10} \cdot \varepsilon$. Similarly, if a red set $Z \in \mathcal{Z}^4$ intersects a side Γ_j of $\partial[0,1]^2$, then $\lambda_1(Z \cap \Gamma_j \setminus S) \geq \frac{1}{4} 2^{-10} \cdot \varepsilon$.

We consider first the case $|\mathcal{Z}^4| = 1$. If Z_1^4 is green, then by taking $H \subset \mathbb{R}^2$ to be a halfplane with $[0,1]^2 \subset H$, we see that S is of type (II). Similarly, if Z_1^4 is red, we conclude again that S is of type (II), by taking $H \subset \mathbb{R}^2$ to be a halfplane with $H \cap [0,1]^2 = \emptyset$.

Next, consider the case $|\mathcal{Z}^4| = 2$. If Z_1^4 is green and Z_2^4 is red, then let $H \subset \mathbb{R}^2$ be the halfplane with $H \cap \partial[0,1]^2 = Z_1^4$. Otherwise, if Z_1^4 is red, and Z_2^4 is green then let $H \subset \mathbb{R}^2$ be the halfplane with $H \cap \partial[0,1]^2 = Z_2^4$. It follows in either case that S is of type (II).

It remains to consider the case $|\mathcal{Z}^4| \geq 4$. Assume w.l.o.g. that for any odd i, the set Z_i^4 is green, and for every even i, the set Z_i^4 is red (the remaining case is symmetric). Suppose first that there exists $j \in \{0, \ldots, 3\}$, such that for any odd $i \in \{1, \ldots, |\mathcal{Z}^4|\}$, we have $Z_i \subseteq \Gamma_j$. It follows that S is of type (I). Similarly, if there exists $j \in \{0, \ldots, 3\}$, such that for any even $i \in \{1, \ldots, |\mathcal{Z}^4|\}$, we have $Z_i \subseteq \Gamma_j$, then it again follows that S is of type (I).

It suffices therefore to consider the case where there is no side of $\partial[0,1]^2$ containing either $Y = \bigcup_{i:i \text{ even}} Z_i^4$, or $Y' = \bigcup_{i:i \text{ odd}} Z_i^4$. It follows that there exist $Y_1, Y_3 \subset Y$, and $Y_2, Y_4 \subset Y'$, satisfying the following conditions:

- For any $i \in \{1,3\}$, we have $\lambda_1(Y_i \cap S) \geq 2^{-11} \cdot \varepsilon$.
- For any $i \in \{2,4\}$, we have $\lambda_1(Y_i \setminus S) \geq 2^{-11} \cdot \varepsilon$.
- The sets Y_1, Y_2, Y_3, Y_4 appear in the order they are listed along a clock-wise traversal of $\partial[0,1]^2$.
- For every $i \in \{1, \ldots, 4\}$, there exists $j \in \{1, \ldots, 4\}$, such that $Y_i \subseteq \Gamma_j$.
- Y_1 and Y_2 are contained in different sides of $\partial[0,1]^2$.
- Y_3 and Y_4 are contained in different sides of $\partial[0,1]^2$.

Setting for every $i \in \{1,3\}$, $I_i = Y_i \cap S$, and for any $i \in \{2,4\}$, $I_i = Y_i \setminus S$, we obtain sets $I_1, \ldots, I_4$, satisfying the conditions of Lemma C.5. It follows by Lemma C.5 that S is not of ε-low complexity, a contradiction. We have obtained that in all cases, S is either of type (I), or of type (II), as required. $\square$

Lemma C.9 (Type (II) sets)**.** *Let $S \subseteq [0,1]^2$ be a ε-low complexity set of type (II). Then, there exists a halfplane $H \subset \mathbb{R}^2$, such that if we set $T = H \cap [0,1]^2$, then $\lambda_2(S \triangle T) \leq \varepsilon$. Moreover, we can choose such a halfplane H, depending only on the set $S \cap \partial[0,1]^2$.*

Proof. By the definition of a type (II) set, we have that there exists a halfplane $H \subset \mathbb{R}^2$, such that

$$\lambda_1(\partial[0,1]^2 \cap H \cap S) \geq \lambda_1(\partial[0,1]^2 \cap H) - 2^{-7} \cdot \varepsilon,$$

and

$$\lambda_1((\partial[0,1]^2 \setminus H) \setminus S) \geq \lambda_1(\partial[0,1]^2 \setminus H) - 2^{-7} \cdot \varepsilon.$$

Suppose for the sake of contradiction that $\lambda_2(S \triangle T) > \varepsilon$. W.l.o.g., we can assume that $\lambda_2(S \setminus T) > \varepsilon/2$, since the remaining case where $\lambda_2(([0,1]^2 \setminus S) \setminus ([0,1]^2 \setminus T)) > \varepsilon/2$, can be treated analogously by replacing S with $[0,1]^2 \setminus S$.

Consider the halfplane $H' = \{x \in \mathbb{R}^2 : \exists y \in H, \|x - y\| \leq \varepsilon/8\}$. Let $S' = S \setminus H'$. We have $\lambda_2(S') > \varepsilon/4$.

Let θ be the angle of the line supporting H. Let $\phi \in [\theta - \varepsilon/32, \theta + \varepsilon/32]$. Let

$$\Lambda'_\phi = \{\ell \in \Lambda_\phi : \ell \cap T = \emptyset\}.$$

We have

$$\int_{\Lambda_{\phi'}} \lambda_1(\ell \cap S') d\mu_\phi(\ell) \geq \lambda_2(S') \geq \varepsilon/4. \tag{15}$$

Let

$$\Lambda''_\phi = \{\ell \in \Lambda'_\phi : \lambda_1(\ell \cap S) \geq \varepsilon/8\}.$$

Since for every $\ell \in \Lambda$, we have $\lambda_1(\ell \cap S) \leq \sqrt{2}$, it follows by (15) that

$$\mu_\phi\left(\Lambda''_\phi\right) > \frac{\varepsilon}{32}.$$

Let

$$\Lambda'''_\phi = \{\ell \in \Lambda''_\phi : \ell \cap \partial[0,1]^2 \cap S = \emptyset\}.$$

We have

$$\mu_\phi\left(\Lambda'''_\phi\right) \geq \mu_\phi\left(\Lambda''_\phi\right) - \sqrt{2} \cdot 2^{-7} \cdot \varepsilon > \frac{\varepsilon}{64}. \tag{16}$$

By Lemma C.3 and (16) we obtain

$$|\mathcal{C}_\eta(S) - \mathcal{C}(S)| > \frac{\varepsilon}{16} \cdot \frac{\varepsilon}{64} \cdot \frac{1}{2} = \frac{\varepsilon^2}{2048} > 2^{-27} \cdot \varepsilon^2,$$

contradicting the fact that S is of ε-low complexity, and concluding the proof. $\square$

Lemma C.10 (Type (I) sets)**.** *Let $S \subseteq [0,1]^2$ be a ε-low complexity set. If S is of type (I-a), then $\lambda_2(S) \geq 1 - \varepsilon$, and if S is of type (I-b), then $\lambda_2(S) \leq \varepsilon$.*

Proof. The proof is identical to the proof of Lemma C.9 for sets of type (II). In particular, if S is of type (I-a), and $|(\Gamma_i \cup \Gamma_{i+1} \cup \Gamma_{i+2}) \cap S| \geq |\Gamma_i \cup \Gamma_{i+1} \cup \Gamma_{i+2}| - 2^{-7} \cdot \varepsilon$, for some $i \in \{1, \ldots, 4\}$, then we set H to be the halfplane with supporting line containing Γ_{i+3}, and such that $[0,1]^2 \subset H$. Otherwise, if S is of type (I-b), and $|(\Gamma_i \cup \Gamma_{i+1} \cup \Gamma_{i+2}) \setminus S| \geq |\Gamma_i \cup \Gamma_{i+1} \cup \Gamma_{i+2}| - 2^{-7} \cdot \varepsilon$, for some $i \in \{1, \ldots, 4\}$, then we set H to be the halfplane with supporting line containing Γ_{i+3}, and such that $[0,1]^2 \cap H = \Gamma_{i+3}$. The rest of the argument is exactly the same as in Lemma C.9. $\square$

C.2 Pairs of ε-low complexity sets

Lemma C.11. *Let $\varepsilon \in \left(0, \frac{1}{64}\right)$, $\eta = \frac{1}{2} \cdot \varepsilon^6$. Let $S_0, S_1 \subseteq [0,1]^2$, with $S_0 \cap \partial[0,1]^2 = S_1 \cap \partial[0,1]^2$. Suppose that both S_0 and S_1 are of ε-low complexity. Then, $\lambda_2(S_0 \triangle S_1) \leq 2 \cdot \varepsilon$.*

Proof. By Lemma C.8 we have that for each $i \in \{1,2\}$, the set S_i is of type (I), or of type (II), or both. Observe that S_0 is of type (I-a), (I-b), or (II), if and only if S_1 is also of type (I-a), (I-b), or (II) respectively.

If both S_0 and S_1 are of type (I-a), then by Lemma C.10 we have

$$\lambda_2(S_0 \triangle S_1) \leq \lambda_2([0,1]^2 \setminus S_0) + \lambda_2([0,1]^2 \setminus S_1) \leq 2 \cdot \varepsilon.$$

Otherwise, if they are both of type (I-b), then again by Lemma C.10 we have

$$\lambda_2(S_0 \triangle S_1) \leq \lambda_2(S_0) + \lambda_2(S_1) \leq 2 \cdot \varepsilon.$$

Finally, consider the case where both S_0 and S_1 are of type (II). By Lemma C.9 there exist halfplanes $H_0, H_1 \subset \mathbb{R}^2$, so that if we set $T_0 = H_0 \cap [0,1]^2$, and $T_1 = H_1 \cap [0,1]^2$, we have for any $i = \{0,1\}$,

$$\lambda_2(S_i \triangle T_i) \leq \varepsilon.$$

However, by Lemma C.9 it follows that for each $i \in \{0,1\}$, the choice of H_i depends only on $S_i \cap \partial[0,1]^2$. Since $S_0 \cap \partial[0,1]^2 = S_1 \cap \partial[0,1]^2$, it follows that we can choose $H_0 = H_1$, and thus $T_0 = T_1$. We conclude that

$$\lambda_2(S_0 \triangle S_1) \leq \lambda_2(S_0 \triangle T_0) + \lambda_2(S_1 \triangle T_1) \leq 2 \cdot \varepsilon,$$

as required. □

C.3 Inner portions of small sets of ε-low complexity

We will use the following isoperimetric property of subsets of $\mathbb{R}^2$. We remark that if we require that the inner probability is at least $\Omega(\varepsilon)$, then asymptotically the bound is best possible, e.g. for the set $X = [0, \varepsilon^2] \times [0,1]$. Moreover, one cannot require that the inner probability is $\omega(\varepsilon)$, as can be seen by setting $X = [0, \varepsilon]^2$.

Lemma C.12 (Intersecting sets with random lines)**.** *Let $X \subseteq [0,1]^2$ be a polygonal set, with $\lambda_2(X) \geq \varepsilon^2$. Then,*

$$\mathbb{P}_{\theta \in [0,2\pi)} \left[\mathbb{P}_{\ell \in \Lambda_\theta} \left[\lambda_1(\ell \cap X) \geq \frac{1}{4 \cdot \sqrt{2}} \cdot \varepsilon^2 \right] \geq \frac{1}{4} \cdot \varepsilon \right] \geq \frac{1}{2}.$$

Proof. For every $\theta \in [0, 2\pi)$, let

$$A_\theta = \left\{ \ell \in \Lambda_\theta : \lambda_1(\ell \cap X) \geq \frac{1}{4 \cdot \sqrt{2}} \cdot \varepsilon^2 \right\}.$$

If for all $\theta \in [0,2\pi)$, we have $\mu_\theta(A_\theta) \geq \frac{\sqrt{2}}{4} \cdot \varepsilon$, then the assertion clearly holds. So, assume w.l.o.g. for the remainder of the proof that there exists $\theta^* \in [0, 2\pi)$, such that

$$\mu_{\theta^*}(A_{\theta^*}) < \frac{\sqrt{2}}{4} \cdot \varepsilon. \tag{17}$$

Let

$$S' = \bigcup_{\ell \in A_{\theta^*}} (S \cap \ell).$$

We have

$$\lambda_2(S \setminus S') < \frac{1}{4} \cdot \varepsilon^2. \tag{18}$$

Since $S' \subseteq S$, it follows by (18) that

$$\lambda_2(S') = \lambda_2(S) - \lambda_2(S \setminus S') > \frac{1}{2} \cdot \varepsilon^2. \tag{19}$$

Observe that by (17), for any $\ell' \in \Lambda_{\theta^*+\pi/2}$, we have

$$\lambda_1(\ell' \cap S') \leq \mu_{\theta^*}(A_{\theta^*}) < \frac{\sqrt{2}}{4} \cdot \varepsilon.$$

Since

$$\lambda_2(S') = \int_{\Lambda_{\theta^*+\pi/2}} \lambda_1(\ell' \cap S') d\ell',$$

it follows that

$$\mu_{\theta^*+\pi/2}(A_{\theta^*+\pi/2}) > \frac{\sqrt{2}}{4} \cdot \varepsilon.$$

We have obtained that for every $\theta^* \in [0, 2\pi)$, such that $\mu_{\theta^*}(A_{\theta^*}) < \frac{\sqrt{2}}{4} \cdot \varepsilon$, we have $\mu_{\theta^*+\pi/2}(A_{\theta^*+\pi/2}) > \frac{\sqrt{2}}{4} \cdot \varepsilon$. Let

$$\Phi = \left\{ \theta \in [0, 2\pi) : \mu_\theta(A_\theta) < \frac{\sqrt{2}}{4} \cdot \varepsilon \right\},$$

and

$$\Psi = \left\{ \theta \in [0, 2\pi) : \mu_\theta(A_\theta) \geq \frac{\sqrt{2}}{4} \cdot \varepsilon \right\}.$$

Let μ denote the standard measure on $[0, 2\pi)$. It follows that $\Phi + \pi/2 = \{\theta + \pi/2 : \theta \in \Phi\} \subseteq \Psi$, which implies that $\mu(\Psi) \geq \mu(\Phi)$, and therefore $\mu(\Phi) \geq \pi$, concluding the proof. $\square$

Lemma C.13 (Inner portions of small sets)**.** *Let $\varepsilon \in \left(0, \frac{1}{64}\right)$, $\eta = \frac{1}{2} \cdot \varepsilon^6$. Let $S \subset [0,1]^2$ be of ε-low complexity, with $\lambda_2(S) \leq 1/64$. Then, $\lambda_2(inner(S)) \leq \varepsilon^2$.*

Proof. Since S is of ε-low complexity, it follows by Lemma C.8 that S is of type (I), or of type (II) (or both). Since $\lambda_2(S) \leq 1/64 < 1 - \varepsilon$, it follows by Lemma C.10 that S cannot be of type (I-a).

We can therefore assume w.l.o.g. that S is of type (I-b), or of type (II). In either case, it follows that there exists a halfplane $H \subset \mathbb{R}^2$, such that

$$\lambda_1((\partial[0,1]^2 \setminus H) \setminus S) \geq \lambda_1(\partial[0,1]^2 \setminus H) - 2^{-7} \cdot \varepsilon. \tag{20}$$

Let $T = H \cap [0,1]^2$. By Lemmas C.9, and C.10 we have

$$\lambda_2(S \triangle T) \leq \varepsilon.$$

Thus,

$$\lambda_2(T) \leq \frac{1}{64} + \varepsilon \leq \frac{1}{32}. \tag{21}$$

Let

$$A = \left\{ \theta \in [0, 2\pi) : \mathbb{P}_{\ell \in \Lambda_\theta} \left[\lambda_1(\ell \cap \operatorname{inner}(S)) \geq \frac{1}{4 \cdot \sqrt{2}} \cdot \varepsilon^2 \right] \geq \frac{1}{4} \cdot \varepsilon \right\}.$$

Let μ denote the standard measure on $[0, 2\pi)$. We have by Lemma C.12, we have

$$\mu(A) \geq \pi. \tag{22}$$

Let

$$A' = \{\theta \in A : \forall \ell \in \Lambda_\theta, \text{ if } \ell \cap \operatorname{inner}([0,1]^2) \neq \emptyset, \text{ then } \ell \cap T = \emptyset\}.$$

By (21) and (22), it follows that

$$\mu(A') \geq \pi/4. \tag{23}$$

For any $\theta \in A'$, let

$$A'_\theta = \left\{ \ell \in \Lambda_\theta : \ell_1(\ell \cap S) \geq \frac{1}{4 \cdot \sqrt{2}} \cdot \varepsilon^2, \text{ and } \ell \cap (\partial[0,1]^2) \cap S = \emptyset \right\}.$$

By (20) we have that for any $\theta \in A'$,

$$\mu_\theta(A'_\theta) \geq \frac{\sqrt{2}}{4} \cdot \varepsilon - 2^{-17} \cdot \varepsilon > \frac{\sqrt{2}}{8} \cdot \varepsilon. \tag{24}$$

Combining (23), (24), and Lemma C.3, we obtain

$$|\mathcal{C}_\eta(S) - \mathcal{C}(S)| \geq \frac{1}{2^7} \cdot \varepsilon > 2^{-27} \cdot \varepsilon^2,$$

contradicting the fact that S is of ε-low complexity. We remark that since we have $\eta = \Theta(\varepsilon^6)$, it would have been sufficient for Lemma C.3 to consider in the above argument lines ℓ with $\lambda_1(\ell \cap S) \geq \varepsilon^6$, instead of ε^2. This however does not affect the overall bound. This concludes the proof. $\square$