

# Near-optimal Distortion Bounds for Embedding Doubling Spaces into $L_1$

[Extended Abstract]

James R. Lee<sup>\*</sup>  
University of Washington

Anastasios Sidiropoulos  
Toyota Technology Institute

## ABSTRACT

We exhibit an infinite doubling metric space  $(\mathcal{X}, d)$  such that for any non-expansive mapping  $f : \mathcal{X} \rightarrow L_1$ , 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].

## Categories and Subject Descriptors

F.2.2 [Analysis of Algorithms and Problem Complexity]: Nonnumerical Algorithms and Problems

## General Terms

Algorithms, Theory

## 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

<sup>\*</sup>Supported by NSF grants CCF-0915251 and CCF-0644037, and a Sloan Research Fellowship.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

STOC'11, June 6-8, 2011, San Jose, California, USA.

Copyright 2011 ACM 978-1-4503-0691-1/11/06 ...\$10.00.

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  $\text{cap}, \text{dem} : V \times V \rightarrow \mathbb{R}_{\geq 0}$ , one defines the *sparsity* of the subset  $S \subseteq V$  by

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

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

It was shown in [LLR95, AR98, GNRS99] that the integrality gap for a natural 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:

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

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  $f : X \rightarrow L_2$  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  $c_1(X, d) \leq O(1)$  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 [CKN10] 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  $c_2(X, \sqrt{d}) \leq D$ . 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-Carathéodory 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.

**Doubling metric spaces.** 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 [CKN10] 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 [CKN10] 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; in order to obtain the correct asymptotic dependence, 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 [CKN10]. 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.

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

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

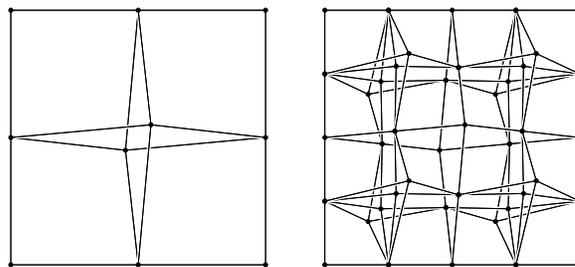
for some constant  $C = O(1)$ , independent of the sequence. This is merely the weak negative type condition in disguise: It simply says that  $W$  is the image of a weak negative type embedding of some metric space. In all known algorithmic applications, it is only the weak condition that is needed.

Far more than being a curiosity 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  $\tilde{O}(n^2)$  time. In [She09], such an approximation is obtained in  $\tilde{O}(m + n^{3/2+\varepsilon})$ -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  $f : \mathbb{R} \rightarrow L_1$  are not guaranteed to be differentiable in the classical sense.

More recently, 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, CKN10] 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  $f : \mathbb{R} \rightarrow \mathbb{R}$  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



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

mappings  $f : [0, 1]^2 \rightarrow L_1$  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. In order to obtain a quantitatively near-optimal result, we achieve a tight quantitative form of this classification for subsets of  $[0, 1]^2$ . Our quantitative analysis follows roughly along the same lines as that of [CKN10], with a number of changes adapted to our particular setting, and the goal of an essentially optimal analysis.

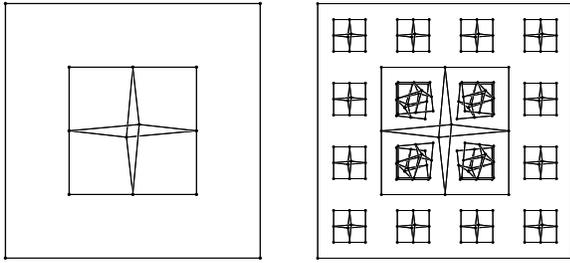
A major difficulty is that since we must work at a fixed scale, we do not have monotonicity with respect to lines, but instead with respect to discrete sequences of points along these lines. If these points were deterministic, then even very bad sets could elude our test lines by being periodic in sync with the discrete sequence. To counter this, we develop a “random” discrete differentiation theory. For instance, consider a mapping  $f : [0, 1] \rightarrow L_1$ . Instead of subdividing  $[0, 1]$  into a sequence of hierarchical partitions, we use a sequence of random progressively finer, *non-hierarchical* subdivisions. We remark that in the approach of [CKN10, §6], a related issue is handled using the kinematic formula for Carnot groups.

## 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 novel aspects that go into proving a precise quantitative bound.

**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  $\partial S$  to denote the topological boundary of a set  $S \subseteq \mathbb{R}^2$ . 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



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

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.

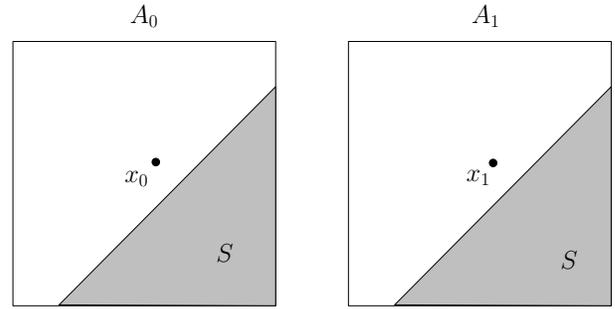
The diamondfolds are *not* doubling, but they are slightly easier to reason about than their doubling counterparts, the *Laakso-folds*, which we denote by  $\{\mathcal{L}_k\}$ . The Laakso-folds are based on the Laakso graphs [Laa02, LP01, GKL03]. These are constructed in the same manner as the diamondfold,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

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 (i.e., weighted graph) 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 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  $k = 0, 1, \dots$  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.*



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

## 2.1 A qualitative lower bound

We first give an intuitive (and ultimately invalid) sketch of why  $c_1(\mathcal{X}_k) \rightarrow \infty$ . 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  $f : X \rightarrow L_1$ . It is well-known that there exists a measure  $\nu$  on  $2^X$ , the subsets of  $X$ , such that for any  $x, y \in X$ ,

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

Here,  $\mathbf{1}_S$  denotes the characteristic function of  $S$ .

Now, examine a Lipschitz mapping  $f : [0, 1]^2 \rightarrow L_1$ , and let  $\nu$  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  $f|_\ell : \ell \rightarrow L_1$  and the corresponding cut measure  $\nu_\ell$  on  $\ell$ : By a differentiation argument, one can show, morally, that on small subintervals  $[a, b] \subseteq \ell$ , the cut measure  $\nu_\ell$  restricted to  $[a, b]$  is concentrated almost entirely on half-segments of the form  $[a, b']$  for some  $b' \leq b$ . This “monotonicity” 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).

When this information is combined from all lines  $\ell$ , 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  $\nu_A$  is the cut measure restricted to  $A$ , then  $\nu_A$  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  $f : [0, 1]^2 \rightarrow L_1$ , there exists a square  $A \in \mathcal{D}_k$ , where the restricted cut measure  $\nu_A$  is concentrated *entirely* on halfspace cuts. In this case, given a 1-Lipschitz mapping  $F : \mathcal{X}_k \rightarrow L_1$ , 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  $\nu_{A_0}$  and  $\nu_{A_1}$  are both concentrated entirely on halfspace cuts. Otherwise, we could apply Proposition 2.1 and conclude that there is an isometric copy of  $C$   $[0, 1]^2$  where all  $A \in \mathcal{D}_k$  have some of their cut measure *not* on halfspace cuts, and this would contradict our assumption for the map  $F|_C : [0, 1]^2 \rightarrow L_1$ .

Now, take these sheets  $A_0$  and  $A_1$ , and let  $\nu_S$  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  $\nu_S$ -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  $\varepsilon$ ’s and  $\delta$ ’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.

## 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).

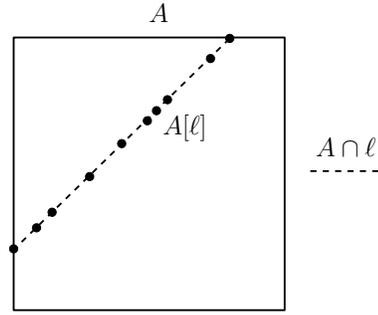
**Setup: The kinematic measure and randomly sprinkled needles.** Let  $\mu_0$  be the kinematic measure on lines  $\ell \subseteq \mathbb{R}^2$ . This is the unique measure, up to scaling, which is invariant under rigid motions. Let  $\Lambda$  be the set of all lines in  $\mathbb{R}^2$  that intersect  $[0, 1]^2$ , 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 metric space. For a function  $F : [0, 1]^2 \rightarrow Y$  and any finite sequence  $s = \langle s_1, s_2, \dots, s_j \rangle$  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  $s$ . We prove the following result. (For reasons of space, the reader is referred to the full version of the paper.)



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

**THEOREM 2.2.** *For every metric space  $Y$  and 1-Lipschitz mapping  $F : [0, 1]^2 \rightarrow Y$ , 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*

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

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

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

(Here, we have used the fact that the  $\mu$ -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  $\ell$  and a random set  $A[\ell]$ , then the expected variation of  $F$  over 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  $F : [0, 1]^2 \rightarrow L_1$ , and let  $\nu$  be the corresponding cut measure on  $A$ , then after an application of Fubini’s theorem, the conclusion is that

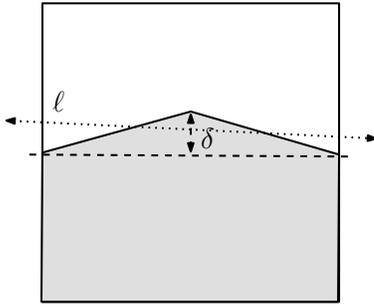
$$\int \int \mathbb{E}[\Delta_{\mathbf{1}_S}(A[\ell]) - \Delta_{\mathbf{1}_S}(\partial A \cap \ell)] d\mu(\ell) d\nu(S) \leq O(\varepsilon). \quad (3)$$

For a set  $S$ , let

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

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  $\ell$  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,  $\theta_S^k = 0$  for a halfspace  $H$ . We know from (3) that  $\int \theta_S^k d\nu(S) \leq O(\varepsilon)$ . Hence our next goal, the “classification” step, is to show the reverse implication: if  $\theta_S$  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



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

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. We use  $\lambda_2$  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 polygonal set. Then there exists a halfspace  $H$  such that*

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

In other words, sets with  $\theta_S^k$  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.

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

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

To finish the argument, we need one additional theorem which relies on the classification lemma.

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

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

**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  $\tilde{N} \subseteq \mathcal{X}_r$  be a  $\delta$ -net, for some  $\delta$  to be specified later. Note that  $|\tilde{N}| \leq O(2^r \delta^{-2})$ .

Suppose that  $F_0 : N_\delta \rightarrow L_1$  is a 1-Lipschitz mapping with distortion  $D$ . A simple Lipschitz extension argument shows

that there is an  $O(1)$ -Lipschitz extension  $F : \mathcal{X}_r \rightarrow L_1$ . 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  $\delta \geq 2^{-r}/O(D)$ , 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

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

where  $\pi_0, \pi_1$  are the natural maps from  $[0, 1]^2$  to  $A_0$  and  $A_1$ , respectively. This follows from  $d_{\mathcal{X}_r}(\pi_0(x), \pi_1(x)) \gtrsim \text{diam}(A_0)$  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  $\mathbb{S}$  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  $\nu$  be the cut measure induced by  $F$ , restricted to  $\mathbb{S}$ , we have

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

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  $\nu_{\text{inner}}$  for the measure  $\nu$  restricted to sets  $S$  which satisfy

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

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  $1/64$  ( $\nu_{\text{large}}$ ), both sets have measure at most  $1/64$ , but one set has measure greater than  $\varepsilon$  ( $\nu_{\text{small}}$ ), both sets have measure at most  $\varepsilon$ .

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

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

We bound each of these three terms separately.

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

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

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

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

by (5), and the fact that

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

The latter fact follows because  $\nu_{\text{large}}$  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  $\int d\nu_{\text{large}} = O(1)$ .

We use Theorem 2.5 to bound,

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

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

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

$$\begin{aligned} \int \rho(S) d\nu_{\text{tiny}}(S) &\leq \int \lambda_2(S \cap A_0) + \lambda_2(S \cap A_1) d\nu_{\text{tiny}}(S) \\ &\lesssim \varepsilon \cdot \frac{1}{\sqrt{\varepsilon}} = \sqrt{\varepsilon}. \end{aligned}$$

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  $F : [0, 1]^2 \rightarrow Y$ , 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*

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

### 3. 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 Symposium 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.
- [CKN10] J. Cheeger, B. Kleiner, and A. Naor. Compression bounds for Lipschitz maps from the Heisenberg group to  $L_1$ . arXiv:0910.2026. To appear, *Acta Math.*, 2010.
- [CMM06] Eden Chlamtac, Konstantin Makarychev, and Yury Makarychev. How to play unique game using embeddings. In *47th Annual Symposium 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  $\ell_1$ . In *Proceedings of the 46th Annual IEEE Conference on Foundations of Computer Science*, pages 53–62, 2005.
- [Laa02] Tomi J. Laakso. Plane with  $A_\infty$ -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. Raghavendra. 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.