Metric structures in L_1: Dimension, snowflakes, and average distortion

James R. Lee
U.C. Berkeley and Microsoft Research
jrl@cs.berkeley.edu

Manor Mendel
Hebrew University
mendelma@cs.huji.ac.il

Assaf Naor
Microsoft Research
anaor@microsoft.com

Abstract

We study the metric properties of finite subsets of L_1. The analysis of such metrics is central to a number of important algorithmic problems involving the cut structure of weighted graphs, including the Sparsest Cut Problem, one of the most compelling open problems in the field of approximation. We present some new observations concerning the relation of L_1 to dimension, topology, and Euclidean distortion. In particular, we offer new insights into the four main open problems surrounding the metric structure of L_1.

1 Introduction

This paper is devoted to the analysis of metric properties of finite subsets of L_1. Such metrics occur in many important algorithmic contexts, and their analysis is key to progress on some fundamental problems. For instance, an $O(\log n)$-approximate max-flow/min-cut theorem proved elusive for many years until, in [15, 2], it was shown to follow from a theorem of Bourgain stating that every metric on n points embeds into L_1 with distortion $O(\log n)$.

The importance of L_1 metrics has given rise to many problems and conjectures that have attracted a lot of attention in recent years. Four basic problems of this type are as follows.

I. Is there an L_1 analog of the Johnson-Lindenstrauss dimension reduction lemma [9]?

II. Are all n-point subsets of L_1 $O\left(\sqrt{\log n}\right)$-embeddable into Hilbert space?

III. Are all squared-ℓ_2 metrics $O(1)$-embeddable into L_1?

IV. Are all planar graphs $O(1)$-embeddable into L_1?

Each of these questions has been asked many times before; we refer to [18, 19, 14, 8], in particular. Despite an immense amount of interest and effort, the metric properties of L_1 have proved quite elusive; hence the name "The mysterious L_1" appearing in a survey of Linial at the ICM in 2002 [14]. In this paper, we offer new insights into each of the above problems and touch on some relationships between them.

1.1 Results and techniques

Dimension reduction. In [3], and soon after in [13], it was shown that if the Newman-Rabinovich diamond graph on n vertices α-embeds into ℓ_1^d then $d \geq n^{\Omega(1/\alpha^2)}$. The proof in [3] is based on a linear programming argument, while the proof in [13] uses a geometric argument which reduces the problem to bounding from below the distortion required to embed the diamond graph in ℓ_p, $1 < p < 2$. These results settle the long standing open problem whether there is an L_1 analog of the Johnson-Lindenstrauss dimension reduction lemma [9]. (In other words, they show that the answer to problem (I) above is *No.*). In Section 2, we show that the method of proof in [13] can be used to provide an even more striking counter example to this problem.

A metric space X is called *doubling* with constant C if every ball in X can be covered by C balls of half the radius. Doubling metrics with bounded doubling constants are widely viewed as low dimensional (see [6, 10] for some practical and theoretical applications of this viewpoint). In fact, they have bounded Assouad dimension (see [7] for the definition). On the other hand, the doubling constant of the diamond graphs is $\Omega(\sqrt{n})$ (where n is the number of points). Based on a fractal construction due to Laakso [11] and the method developed in [13], we prove the following theorem, which shows a strong lower bound on the dimension required to represent uniformly doubling subsets of L_1.

Theorem 1.1. *There are arbitrarily large n-point subsets $X \subseteq L_1$ which are doubling with constant 6 but such that every α-embedding of X into ℓ_1^d requires $d \geq n^{\Omega(1/\alpha^2)}$.*

In [12, 6] it was asked whether any subset of ℓ_2 which is doubling well-embeds into ℓ_2^d (with bounds on the distortion and the dimension that depend only on the doubling constant). In [6], it was shown that a similar property cannot hold for ℓ_1. Our lower bound exponentially strengthens this result.

Planar metrics. The next result addresses problems (III) and (IV). Our motivation was an attempt to generalize the argument in [13] to prove that dimension reduction is impossible in L_p for any $1 < p < 2$. A natural approach to this problem is to consider the point set used in [3, 13] (namely, a natural realization of the diamond graph, G, in L_1) with the metric induced by the L_p norm instead of the L_1 norm. This is easily seen to amount to proving lower bounds on the dimension required to embed the metric space $(G, d_G^{1/p})$ in ℓ_p^d. Unfortunately, this approach cannot work since we show that, for any planar metric (X, d) and any $0 < \varepsilon < 1$, the metric space $(X, d^{1-\varepsilon})$ embeds in Hilbert space with distortion $O\left(1/\sqrt{\varepsilon}\right)$. The proof of this interesting fact is a straightforward application of Assouad's classical embedding theorem [1] and Rao's embedding method [23]. The $O\left(1/\sqrt{\varepsilon}\right)$ upper bound is shown to be tight for every value $0 < \varepsilon < 1$. The case $\varepsilon = 1/2$ has been previously observed by A. Gupta in his (unpublished) thesis. It follows that any planar metric embeds into squared-ℓ_2 with $O(1)$ distortion so that a positive solution to problem (III) above implies a positive solution to problem (IV).

Euclidean distortion. Our final result addresses problem (II) stated above (and for more subtle reasons, problem (III)). We show that the answer to this question is positive on average, in the following sense.

Theorem 1.2. *For every $f_1, \ldots, f_n \in L_1$ there is a linear operator $T : L_1 \to L_2$ such that*

$$\frac{1}{(8\log n)^{1/3}} \le \min_{1\le i<j\le n} \left(\frac{\|T(f_i) - T(f_j)\|_2}{\|f_i - f_j\|_1}\right)^{2/3} \le \frac{1}{\binom{n}{2}} \sum_{1\le i<j\le n} \left(\frac{\|T(f_i) - T(f_j)\|_2}{\|f_i - f_j\|_1}\right)^{2/3} \le 10.$$

In other words, for any n-point subset in L_1, there exists a map into L_2 such that distances are contracted by at most $O(\sqrt{\log n})$ and the average expansion is $O(1)$. We remark that a different notion of average embedding was recently studied by Rabinovich [22]. In [22] one tries to embed (planar) metrics into the line such that the *average distance* does not change too much.

The exponent $2/3$ above has no significance, and we can actually obtain the same result for any power $1-\varepsilon$, $\varepsilon > 0$ (we refer to Section 4 for details). The proof of Theorem 1.2 follows from the following probabilistic lemma, which is implicit in [16]. We believe that this result is of independent interest.

Lemma 1.3. *There exists a distribution over linear mappings $T : L_1 \to L_2$ such that for every $x \in L_1 \setminus \{0\}$ the random variable $\frac{\|T(x)\|_2}{\|x\|_1}$ has density $\frac{e^{-1/(4x^2)}}{x^2\sqrt{\pi}}$.*

In contrast to Theorem 1.2, we show that problem (II) cannot be resolved positively using linear mappings. Specifically, we show that there are arbitrarily large n-point subsets of L_1 such that any linear embedding of them into L_2 incurs distortion $\Omega(\sqrt{n})$. As a corollary we settle the problem left open by Charikar and Sahai in [4], whether *linear* dimension reduction is possible in L_p, $p \notin \{1, 2\}$. The case $p = 1$ was proved in [4] via linear programming techniques, and it seems impossible to generalize their lower bound to arbitrary L_p. We show that there are arbitrarily large n-point subsets $X \subseteq L_p$ (namely, the same point set used in [4] to handle the case $p = 1$), such that any linear embedding of X into ℓ_p^d incurs distortion $\Omega\left[(n/d)^{|1/p-1/2|}\right]$, thus linear dimension reduction is impossible in any L_p, $p \neq 2$. Additionally, we show that there are arbitrarily large n-point subsets $X \subseteq L_1$ such any linear embedding of X into *any* d-dimensional normed space incurs distortion $\Omega\left(\sqrt{n/d}\right)$. This generalizes the Charikar-Sahai result to arbitrary low dimensional norms.

2 An inherently high-dimensional doubling metric in L_1

This section is devoted to the proof of Theorem 1.1. The case $p = 2$ of the proof below reduces to the argument in [11, 12, 21].

Proof of Theorem 1.1. Consider the Laakso graphs, $\{G_i\}_{i=0}^{\infty}$, which are defined as follows. G_0 is the graph on two vertices with one edge. To construct G_i, take six copies of G_{i-1} and scale their metric by a factor of $\frac{1}{4}$. We glue four of them cyclicly by identifying pairs of endpoints, and attach at two opposite gluing points the remaining two copies. See Figure 1 below.

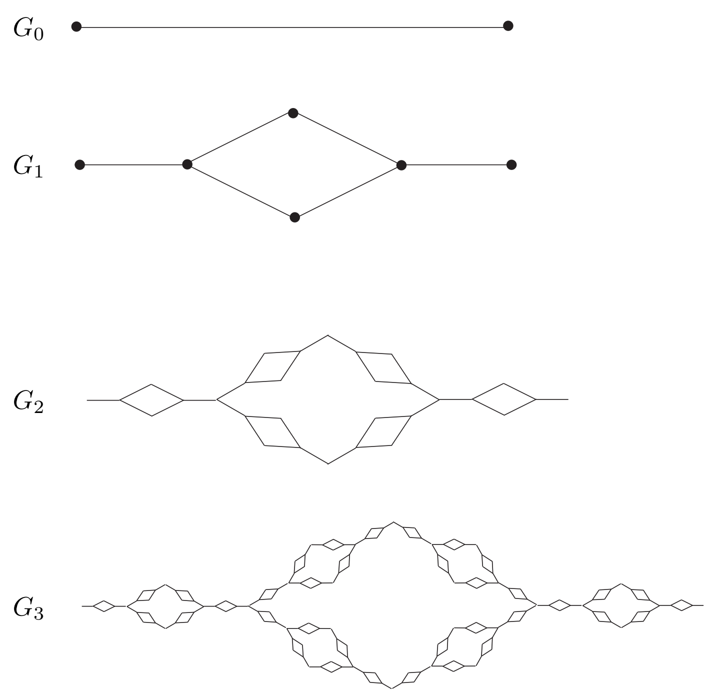

Figure 1: The Laakso graphs.

As shown in [11], the graphs $\{G_i\}_{i=0}^{\infty}$ are uniformly doubling (see also [12], for a simple argument showing they are doubling with constant 6). Moreover, since the G_i's are series parallel graphs, they embed uniformly in L_1 (see [5]).

We will show below that any embedding of G_i in L_p, $1 < p \leq 2$ incurs distortion at least $\sqrt{1 + \frac{p-1}{4} i}$. We then conclude as in [13] by observing that ℓ_1^d is 3-isomorphic to ℓ_p^d when $p = 1 + \frac{1}{\log d}$, so that if G_i embeds with distortion α in ℓ_1^d then $\alpha \geq \sqrt{\frac{i}{40 \log d}}$. This implies the required result since $i \approx \log |G_i|$.

The proof of the lower bound for the distortion required to embed G_i into L_p is by induction on i. We shall prove by induction that whenever $f : G_i \to L_p$ is non-contracting then there exist two adjacent vertices $u, v \in G_i$ such that $\|f(u) - f(v)\|_p \geq d_{G_i}(u, v)\sqrt{1 + \frac{p-1}{4} i}$ (observe that for $u, v \in G_{i-1}$, $d_{G_{i-1}}(u, v) = d_{G_i}(u, v)$). For $i = 0$ there is nothing to prove. For $i \geq 1$, since G_i contains an isometric copy of G_{i-1}, there are $u, v \in G_i$ corresponding to two adjacent vertices in G_{i-1} such that $\|f(u) - f(v)\|_p \geq d_{G_i}(u, v)\sqrt{1 + \frac{p-1}{4}(i-1)}$. Let a, b be the two midpoints between u and v in G_i. By Lemma 2.1 in [13],

$$\begin{aligned}\|f(u)-f(v)\|_p^2 &+ (p-1)\|f(a)-f(b)\|_p^2 \\ &\le \|f(u)-f(a)\|_p^2+\|f(a)-f(v)\|_p^2+\|f(v)-f(b)\|_p^2+\|f(b)-f(u)\|_p^2.\end{aligned}$$

Hence:

$$\begin{aligned}&\max\{\|f(u)-f(a)\|_p^2,\|f(a)-f(v)\|_p^2,\|f(v)-f(b)\|_p^2,\|f(b)-f(u)\|_p^2\}\\ \ge\ & \frac{1}{4}\|f(u)-f(v)\|_p^2+\frac{1}{4}(p-1)\|f(a)-f(b)\|_p^2\\ \ge\ & \frac{1}{4}\left(1+\frac{p-1}{4}(i-1)\right)d_{G_i}(u,v)^2+\frac{p-1}{4}d_{G_i}(a,b)^2\\ =\ & \frac{1}{4}\left(1+\frac{p-1}{4}i\right)d_{G_i}(u,v)^2\\ =\ & \left(1+\frac{p-1}{4}i\right)\max\{d_{G_i}(u,a)^2,d_{G_i}(a,v)^2,d_{G_i}(v,b)^2,d_{G_i}(b,u)^2\}.\end{aligned}$$

□

We end this section by observing that the above approach also gives a lower bound on the dimension required to embed expanders in ℓ_∞.

Proposition 2.1. *Let G be an n-point constant degree expander which embeds in ℓ_∞^d with distortion at most α. Then $d \ge n^{\Omega(1/\alpha)}$.*

Proof. By Matoušek's lower bound for the distortion required to embed expanders in ℓ_p [17], any embedding of G into ℓ_p incurs distortion $\Omega\left(\frac{\log n}{p}\right)$. Since ℓ_∞^d is $O(1)$-equivalent to $\ell_{\log d}^d$, we deduce that $\alpha \ge \Omega\left(\frac{\log n}{\log d}\right)$. □

We can also obtain a lower bound on the dimension required to embed the Hamming cube $\{0,1\}^k$ into ℓ_∞. Our proof uses a simple concentration argument. An analogous concentration argument yields an alternative proof of Proposition 2.1.

Proposition 2.2. *Assume that $\{0,1\}^k$ embeds into ℓ_∞^d with distortion α. Then $d \ge 2^{k\Omega(1/\alpha^2)}$.*

Proof. Let $f=(f_1,\dots,f_d):\{0,1\}^k\to\ell_\infty^d$ be a contraction such that for every $u,v\in\{0,1\}^d$, $\|f(u)-f(v)\|_\infty \ge \frac{1}{\alpha}d(u,v)$ (where $d(\cdot,\cdot)$ denotes the Hamming metric). Denote by P the uniform probability measure on $\{0,1\}^k$. Since for every $1\le i\le k$, f_i is 1-Lipschitz, the standard isoperimetric inequality on the hypercube implies that $P\left(|f_i(u)-\mathbb{E}f_i|\ge k/(4\alpha)\right)\le e^{-\Omega(k/\alpha^2)}$. On the other hand, if $u,v\in\{0,1\}^k$ are such that $d(u,v)=k$ then there exist $1\le i\le d$ for which $|f_i(u)-f_i(v)|\ge k/\alpha$, implying that $\max\{|f_i(u)-\mathbb{E}f_i|,|f_i(v)-\mathbb{E}f_i|\}>k/(4\alpha)$. By the union bound it follows that $de^{-\Omega(k/\alpha^2)}\ge 1$, as required. □

3 Snowflake versions of planar metrics

The problem of whether there is an analog of the Johnson-Lindenstrauss dimension reduction lemma in L_p, $1<p<2$, is an interesting one which remains open. In view of the above proof and the proof in [13], a natural point set which is a candidate to demonstrate the impossibility of dimension reduction in L_p is the realization of the diamond graph in ℓ_1 which appears in [3], equipped with the ℓ_p metric. Since this point set consists of $0,1$ vectors, this amounts to considering the diamond graph with its metric raised to the power $\frac{1}{p}$. Unfortunately, this approach cannot work; we show below that any planar graph whose metric is raised to the power $1-\varepsilon$ has Euclidean distortion $O\left(1/\sqrt{\varepsilon}\right)$.

Given a metric space (X,d) and $\varepsilon > 0$, the metric space $(X, d^{1-\varepsilon})$ is known in geometric analysis (see e.g. [7]) as the $1-\varepsilon$ snowflake version of (X,d). Assouad's classical theorem [1] states that any snowflake version of a doubling metric space is bi-Lipschitz equivalent to a subset of some finite dimensional Euclidean space. A quantitative version of this result (with bounds on the distortion and the dimension) was obtained in [6]. The following theorem is proved by combining embedding techniques of Rao [23] and Assouad [1]. A similar analysis is also used in [6]. In what follows we call a metric K_r-excluded if it is the metric on a subset of a weighted graph which does not admit a K_r minor. In particular, planar metrics are all K_5-excluded.

Theorem 3.1. *For any $r \in \mathbb{N}$ there exists a constant $C(r)$ such that for every $0 < \epsilon < 1$, a $1-\varepsilon$ snowflake version of a K_r-excluded metric embeds into ℓ_2 with distortion at most $C(r)/\sqrt{\varepsilon}$.*

Our argument is based on the following lemma, the proof of which is contained in [23].

Lemma 3.2. *For every $r \in \mathbb{N}$ there is a constant $\delta = \delta(r)$ such that for every $\rho > 0$ and every K_r-excluded metric (X,d) there exists a finitely supported probability distribution μ on partitions of X with the following properties:*

1. *For every $P \in \operatorname{supp}(\mu)$, and for every $C \in P$, $\operatorname{diam}(C) \le \rho$.*

2. *For every $x \in X$, $\mathbb{E}_\mu \sum_{C \in P} d(x, X \setminus C) \ge \delta\rho$.*

Observe that the sum under the expectation in (2) above actually consists of only one summand.

Proof of Theorem 3.1. Let X be a K_r-excluded metric. For each $n \in \mathbb{Z}$, we define a map ϕ_n as follows. Let μ_n be the probability distribution on partitions of X from Lemma 3.2 with $\rho = 2^{n/(1-\varepsilon)}$. Fix a partition $P \in \operatorname{supp}(\mu_n)$. For any $\sigma \in \{-1,+1\}^{|P|}$, consider σ to be indexed by $C \in P$ so that σ_C has the obvious meaning. Following Rao [23], define

$$\phi_P(x) = \bigoplus_{\sigma \in \{-1,+1\}^{|P|}} \sqrt{\frac{1}{2^{|P|}}} \sum_{C \in P} \sigma_C \cdot d(x, X \setminus C),$$

and write $\phi_n = \bigoplus_{P \in \operatorname{supp}(\mu_n)} \sqrt{\mu_n(P)}\, \phi_P$ (here the symbol $\oplus$ refers to the concatenation operator).

Now, following Assouad [1], let $\{e_i\}_{i \in \mathbb{Z}}$ be an orthonormal basis of ℓ_2, and set

$$\Phi(x) = \sum_{n \in \mathbb{Z}} 2^{-n\varepsilon/(1-\varepsilon)} \phi_n(x) \otimes e_n$$

Claim 3.3. *For every $n \in \mathbb{Z}$, and $x, y \in X$, we have $\|\phi_n(x) - \phi_n(y)\|_2 \le 2 \cdot \min\left\{d(x,y), 2^{n/(1-\varepsilon)}\right\}$. Additionally, if $d(x,y) > 2^{n/(1-\varepsilon)}$, then $\|\phi_n(x) - \phi_n(y)\|_2 \ge \delta\, 2^{n/(1-\varepsilon)}$.*

Proof. For any partition $P \in \mathrm{supp}(\mu_n)$, let C_x, C_y be the clusters of P containing x and y, respectively. Note that since for every $C \in P$, $\mathrm{diam}(C) \leq 2^{n/(1-\varepsilon)}$, when $d(x,y) > 2^{n/(1-\varepsilon)}$, we have $C_x \neq C_y$. In this case,

$$\begin{aligned} ||\phi_P(x) - \phi_P(y)||_2^2 &= \mathbb{E}_{\sigma \in \{-1,+1\}^{|P|}} |\sigma_{C_x} d(x, X \setminus C_x) - \sigma_{C_y} d(y, X \setminus C_y)|^2 \\ &\geq \frac{d(x, X \setminus C_x)^2 + d(y, X \setminus C_y)^2}{2}. \end{aligned}$$

It follows that

$$\begin{aligned} ||\phi_n(x) - \phi_n(y)||_2^2 &= \mathbb{E}_{\mu_n} ||\phi_P(x) - \phi_P(y)||_2^2 \\ &\geq \frac{\mathbb{E}_{\mu_n} d(x, X \setminus C_x)^2 + \mathbb{E}_{\mu_n} d(y, X \setminus C_y)^2}{2} \geq \left(\delta\, 2^{n/(1-\varepsilon)}\right)^2. \end{aligned}$$

On the other hand, for every $x, y \in X$, since $d(x, X \setminus C_x), d(y, X \setminus C_y) \leq 2^{n/(1-\varepsilon)}$, we have that $||\phi_P(x) - \phi_P(y)||_2 \leq 2 \cdot \min\left\{d(x,y), 2^{n/(1-\varepsilon)}\right\}$, hence $||\phi_n(x) - \phi_n(y)||_2 \leq 2 \cdot \min\left\{d(x,y), 2^{n/(1-\varepsilon)}\right\}$. □

To finish the analysis, let us fix $x, y \in X$ and let m be such that $d(x,y)^{1-\varepsilon} \in \left(2^m, 2^{m+1}\right]$. In this case,

$$\begin{aligned} ||\Phi(x) - \Phi(y)||_2^2 &= \sum_{n \in \mathbb{Z}} 2^{-2n\varepsilon/(1-\varepsilon)} \, \|\phi_n(x) - \phi_n(y)\|_2^2 \\ &\leq 4 \sum_{n<m} 2^{2n} + 4d(x,y)^2 \sum_{n \geq m} 2^{-2n\epsilon/(1-\varepsilon)} \\ &= 2^{2m+1} + 4d(x,y)^2 \frac{2^{-2m\varepsilon/(1-\varepsilon)}}{1 - 2^{-2\varepsilon/(1-\varepsilon)}} \\ &= O\left(1/\varepsilon\right) \cdot d(x,y)^{2(1-\varepsilon)} \end{aligned}$$

On the other hand,

$$\|\Phi(x) - \Phi(y)||_2 \geq 2^{-m\epsilon/(1-\varepsilon)} \|\phi_m(x) - \phi_m(y)\|_2 \geq \delta 2^m \geq \frac{\delta}{2} d(x,y)^{1-\varepsilon}.$$

The proof is complete.

□

Remark 3.4. The $O\left(1/\sqrt{\varepsilon}\right)$ upper bound in Theorem 3.1 is tight. In fact, for $i \approx 1/\varepsilon$, the $1-\varepsilon$ snowflake version of the Laakso graph G_i (presented in Section 2) has Euclidean distortion $\Omega\left(1/\sqrt{\varepsilon}\right)$. To see this, let $f : G_i \to \ell_2$ be any non-contracting embedding of $(G_i, d_{G_i}^{1-\varepsilon})$ into ℓ_2. For $j \leq i$ denote by K_j the Lipschitz constant of the restriction of f to $(G_j, d_{G_i}^{1-\varepsilon})$ (as before, we think of G_j as a subset of G_i). Clearly $K_0 = 1$, and the same reasoning as in the proof of Theorem 1.1 shows that for $j \geq 1$, $K_j^2 \geq \frac{K_{j-1}^2}{4^\varepsilon} + \frac{1}{4}$. This implies that $K_i^2 \geq \frac{1}{4} + \frac{1}{4^\varepsilon} + \ldots + \frac{1}{4^{i\varepsilon}} = \Omega(1/\varepsilon)$, as required.

4 Average distortion Euclidean embedding of subsets of L_1

The heart of our argument is the following lemma which is implicit in [16], and which seems to be of independent interest.

Lemma 4.1. *For every $0 < p \le 2$ there is a probability space (Ω, P) such that for every $\omega \in \Omega$ there is a linear operator $T_\omega : L_p \to L_2$ such that for every $x \in L_p \setminus \{0\}$ the random variable $X = \frac{\|T_\omega(x)\|_2}{\|x\|_p}$ satisfies for every $a \in \mathbb{R}$, $\mathbb{E}e^{-aX^2} = e^{-a^{p/2}}$. In particular, for $p = 1$ the density of X is $\frac{e^{-1/(4x^2)}}{x^2\sqrt{\pi}}$.*

Proof. Consider the following three sequences of random variables, $\{Y_j\}_{j\ge1}$, $\{\theta_j\}_{j\ge1}$, $\{g_j\}_{j\ge1}$, such that each variable is independent of the others. For each $j \ge 1$, Y_j is uniformly distributed on $[0,1]$, g_j is a standard Gaussian and θ_j is an exponential random variable, i.e. for $\lambda \ge 0$, $P(\theta_j > \lambda) = e^{-\lambda}$. Set $\Gamma_j = \theta_1 + \cdots + \theta_j$. By Proposition 1.5. in [16], there is a constant $C = C(p)$ such that if we define for $f \in L_p$

$$V(f) = C \sum_{j\ge1} \frac{g_j}{\Gamma_j^{1/p}} f(Y_j),$$

then $\mathbb{E}e^{iV(f)} = e^{-\|f\|_p^p}$.

Assume that the random variables $\{Y_j\}_{j\ge1}$ and $\{\Gamma_j\}_{j\ge1}$ are defined on a probability space (Ω, P) and that $\{g_j\}_{j\ge1}$ are defined on a probability space (Ω', P'), in which case we use the notation $V(f) = V(f;\omega,\omega')$. Define for $\omega \in \Omega$ a linear operator $T_\omega : L_p \to L_2(\Omega', P')$ by $T_\omega(f) = V(f;\omega,\cdot)$. Since for every fixed $\omega \in \Omega$ the random variable $V(f;\omega,\cdot)$ is Gaussian with variance $\|T_\omega(f)\|_2^2$, for every $a \in \mathbb{R}$, $\mathbb{E}_{P'}e^{iaV(s;\omega,\cdot)} = e^{-a^2\|T_\omega(f)\|_2^2}$. Taking expectation with respect to P we find that, $\mathbb{E}_P e^{-a^2\|T_\omega(f)\|_2^2} = e^{-a^p\|f\|_p^p}$. This implies the required identity. The explicit distribution in the case $p = 1$ follows from the fact that the inverse Laplace transform of $x \mapsto e^{-\sqrt{x}}$ is $y \mapsto \frac{e^{-1/(4y)}}{2\sqrt{\pi y^3}}$ (see for example [24]). □

Theorem 4.2. *For every $f_1, \ldots, f_n \in L_1$ there is a linear operator $T : L_1 \to L_2$ such that:*

$$\frac{1}{(8\log n)^{1/3}} \le \min_{1\le i<j\le n} \left(\frac{\|T(f_i) - T(f_j)\|_2}{\|f_i - f_j\|_1}\right)^{2/3} \le \frac{1}{\binom{n}{2}} \sum_{1\le i<j\le n} \left(\frac{\|T(f_i) - T(f_j)\|_2}{\|f_i - f_j\|_1}\right)^{2/3} \le 10.$$

Proof. Using the notation of lemma 4.1 (in the case $p = 1$) we find that for every $a > 0$, $\mathbb{E}e^{-aX^2} = e^{-\sqrt{a}}$. Hence, for every $a, \varepsilon > 0$ and every $1 < i < j \le n$,

$$P\left(\frac{\|T_\omega(f_i) - T_\omega(f_j)\|_2}{\|f_i - f_j\|_1} \le \varepsilon\right) = P\left(e^{-aX^2} \ge e^{-a\varepsilon^2}\right) \le e^{a\varepsilon^2 - \sqrt{a}}.$$

Choosing $a = \frac{1}{4\varepsilon^4}$ the above upper bound becomes $e^{-1/(4\varepsilon^2)}$. Consider the set

$$A = \bigcap_{1\le i<j\le n} \left\{\frac{\|T_\omega(f_i) - T_\omega(f_j)\|_2}{\|f_i - f_j\|_1} \ge \frac{1}{\sqrt{8\log n}}\right\} \subseteq \Omega.$$

By the union bound, $P(A) > \frac{1}{2}$, so that

$$\frac{1}{P(A)}\mathbb{E}\left[\frac{1}{\binom{n}{2}} \sum_{1\le i<j\le n} \left(\frac{\|T_\omega(f_i) - T_\omega(f_j)\|_2}{\|f_i - f_j\|_1}\right)^{2/3}\right] \le 2\mathbb{E}X^{2/3} = \frac{2}{\sqrt{\pi}} \int_0^\infty x^{2/3} \cdot \frac{e^{-1/(4x^2)}}{x^2} dx < 10.$$

It follows that there exists $\omega \in A$ for which the operator $T = T_\omega$ has the desired properties. □

Remark 4.3. There is nothing special about the choice of the the power $2/3$ in Corollary 4.2. When $p = 1$, $\mathbb{E}X = \infty$ but $\mathbb{E}X^{1-\varepsilon} < \infty$ for every $0 < \varepsilon < 1$, so we may write the above average with the power $1-\varepsilon$ replacing the exponent $2/3$. Obvious generalizations of Corollary 4.2 hold true for every $1 < p < 2$, in which case the average distortion is of order $C(p)(\log n)^{1/p-1/2}$ (and the power can be taken to be 1).

5 The impossibility of *linear* dimension reduction in L_p, $p \neq 2$

The above method cannot yield a $O\left(\sqrt{\log n}\right)$ bound on the Euclidean distortion of n-point subsets of L_1. In fact, there are arbitrarily large n-point subsets of L_1 on which any *linear* embedding into L_2 incurs distortion at least $\sqrt{\frac{n-1}{2}}$. This follows from the following simple lemma:

Lemma 5.1. *For every $1 \le p \le \infty$ there are arbitrarily large n-point subsets of L_p on which any linear embedding into L_2 incurs distortion at least $\left(\frac{n-1}{2}\right)^{|1/p-1/2|}$.*

Proof. Let $w_1, \dots, w_{2^k}$ be the rows of the $2^k \times 2^k$ Walsh matrix. Write $w_i = \sum_{j=1}^{2^k} w_{ij} e_j$ where $e_1, \dots, e_{2^k}$ are the standard unit vectors in $\mathbb{R}^{2^k}$. Consider the set $A = \{0\} \cup \{w_i\}_{i=1}^{2^k} \cup \{e_i\}_{i=1}^{2^k} \subset \ell_p$. Let $T : \ell_p \to L_2$ be any linear operator which is non contracting and L-Lipschitz on A. Assume first of all that $1 \le p < 2$. Then:

$$\begin{aligned} 2^{k(1+2/p)} = \sum_{i=1}^{2^k} \|w_i\|_p^2 \; &\le \; \sum_{i=1}^{2^k} \|Tw_i\|_2^2 = \sum_{i=1}^{2^k} \left\| \sum_{j=1}^{2^k} w_{ij} T(e_j) \right\|_2^2 \\ &= \; \sum_{i=1}^{2^k} \sum_{j=1}^{2^k} \langle w_i, w_j \rangle \langle T(e_i), T(e_j) \rangle = 2^k \sum_{j=1}^{2^k} \|T(e_j)\|_2^2 \le 4^k \cdot L^2, \end{aligned}$$

which implies that $L \ge 2^{k(1/p-1/2)} = \left(\frac{|A|-1}{2}\right)^{1/p-1/2}$. When $p > 2$ apply the same reasoning, with the inequalities reversed.

□

We remark that the above point set was also used by Charikar and Sahai [4] to give a lower bound on *linear* dimension reduction in L_1. Their proof used a linear programming argument, which doesn't seem to be generalizable to the the case of L_p, $p > 1$. Lemma 5.1 formally implies their result (with a significantly simpler proof), and in fact proves the impossibility of linear dimension reduction in any L_p, $p \neq 2$. Indeed, if there were a linear operator which embeds A into ℓ_p^d with distortion D then it would also be a $D \cdot d^{|1/p-1/2|}$ embedding into ℓ_2^d. It follows that $D \ge \left(\frac{|A|-1}{2d}\right)^{|1/p-1/2|}$. Similarly, since by John's theorem (see e.g. [20]) any d-dimensional normed space is $\sqrt{d}$ equivalent to Hilbert space, we deduce that there are arbitrarily large n-point subsets of L_1, any embedding of which into any d-dimensional normed space incurs distortion at least $\sqrt{\frac{n-1}{2d}}$.

References

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

[2] 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.

[3] B. Brinkman and M. Charikar. On the impossibility of dimension reduction in ℓ_1. In *Proceedings of the 44th Annual IEEE Conference on Foundations of Computer Science*. ACM, 2003.

[4] M. Charikar and A. Sahai. Dimension reduction in the ℓ_1 norm. In *Proceedings of the 43rd Annual IEEE Conference on Foundations of Computer Science*. ACM, 2002.

[5] A. Gupta, I. Newman, Y. Rabinovich, and A. Sinclair. Cuts, trees and ℓ_1 embeddings. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science*, 1999.

[6] Anupam Gutpa, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low-distortion embeddings. In *Proceedings of the 44th Annual Symposium on Foundations of Computer Science, to appear*, 2003.

[7] J. Heinonen. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.

[8] P. Indyk. Algorithmic applications of low-distortion geometric embeddings. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 10–33. October 2001.

[9] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

[10] R. Krauthgamer and J. R. Lee. Navigating nets: Simple algorithms for proximity search. Submitted, 2003.

[11] T. J. Laakso. Ahlfors Q-regular spaces with arbitrary $Q > 1$ admitting weak Poincaré inequality. *Geom. Funct. Anal.*, 10(1):111–123, 2000.

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

[13] J. R. Lee and A. Naor. Embedding the diamond graph in L_p and dimension reduction in L_1. *Geom. Funct. Anal.*, to appear.

[14] N. Linial. Finite metric spaces - combinatorics, geometry and algorithms. In *Proceedings of the International Congress of Mathematicians III*, pages 573–586, 2002.

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

[16] M. B. Marcus and G. Pisier. Characterizations of almost surely continuous p-stable random Fourier series and strongly stationary processes. *Acta Math.*, 152(3-4):245–301, 1984.

[17] J. Matoušek. On embedding expanders into l_p spaces. *Israel J. Math.*, 102:189–197, 1997.

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

[19] J. Matoušek. Open problems, workshop on discrete metric spaces and their algorithmic appl ications. Haifa, March 2002.

[20] Vitali D. Milman and Gideon Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[21] I. Newman and Y. Rabinovich. A lower bound on the distortion of embedding planar metrics into euclidean space. *Discrete Computational Geometry*, 29(1):77–81, 2003.

[22] Y. Rabinovich. On average distorsion of embedding metrics into the line and into l_1. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing*. ACM, 2003.

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

[24] David Vernon Widder. *The Laplace Transform*. Princeton Mathematical Series, v. 6. Princeton University Press, Princeton, N. J., 1941.