ISRAEL JOURNAL OF MATHEMATICS **123** (2001), 285–301

# ON DOMINATED $l_1$ METRICS*

BY

JIŘÍ MATOUŠEK

*Department of Applied Mathematics, Charles University*
*Malostranské nám. 25, 118 00 Praha 1, Czech Republic*
*e-mail: matousek@kam.mff.cuni.cz*

AND

YURI RABINOVICH

*CS Department, Haifa University*
*Haifa 31905, Israel*
*e-mail: yuri@cs.haifa.ac.il*

ABSTRACT

We introduce and study a class $l_1^{\mathrm{dom}}(\rho)$ of $l_1$-embeddable metrics corresponding to a given metric $\rho$. This class is defined as the set of all convex combinations of $\rho$-dominated line metrics. Such metrics were implicitly used before in several constuctions of low-distortion embeddings into $l_p$-spaces, such as Bourgain's embedding of an arbitrary metric $\rho$ on $n$ points with $O(\log n)$ distortion. Our main result is that the gap between the distortions of embedding of a finite metric $\rho$ of size $n$ into $l_2$ versus into $l_1^{\mathrm{dom}}(\rho)$ is at most $O(\sqrt{\log n})$, and that this bound is essentially tight. A significant part of the paper is devoted to proving lower bounds on distortion of such embeddings. We also discuss some general properties and concrete examples.

* Research by J. M. supported by Charles University grants No. 158/99 and 159/99. Part of the work by Y. R. was done during his visit at the Charles University in Prague partially supported by these grants, by the grant GAČR 201/99/0242, and by Haifa University grant for Promotion of Research.
Received September 27, 1999

## 1. Introduction

Approximate embeddings of finite metric spaces into $l_p$-spaces were first studied mainly in connection with problems in Functional Analysis (Local Theory of Banach spaces; see, for instance, [6], [7], [14], [4]). Currently they constitute a rich research area with intimate links to Functional Analysis, to classical questions of Combinatorial Optimisation, and to the design of approximate algorithms. In the last area, it offered new paradigms which allowed one to solve long-standing open problems. A partial list of results and applications includes [16, 15, 2, 9, 18, 12].

Let $(X, \rho)$ be a metric space. If $f\colon X \to Z$ is an embedding of $X$ into a normed space $(Z, \|.\|)$, we say that $f$ has **distortion at most** $D$ if we have

$$\frac{1}{D}\|f(x) - f(y)\| \leq \rho(x,y) \leq \|f(x) - f(y)\| \quad \text{for all } x, y \in X$$

(that is, $f$ is non-expanding and contracts each distance by the factor of at most $D$). For $(X, \rho)$ finite, we define $c_p(\rho)$ as the infimum of $D \geq 1$ such that there is an embedding with distortion at most $D$ into $l_p^m$ (for some natural number $m$). Here $l_p^m$ denotes the space $\mathbb{R}^m$ with the norm $\|x\|_p = \left(\sum_{i=1}^m |x_i|^p\right)^{1/p}$. For $(X, \rho)$ infinite, we define $c_p(\rho)$ as the supremum of $c_p(\tau)$, where $\tau$ is $\rho$ restricted to a finite subspace of$X$ (in this way, we avoid discussion of infinite-dimensional $l_p$-spaces).

The most important cases, both theoretically and practically, are $p = 1$ and $p = 2$. Much of the existing theory deals with finding good bounds on $c_1(\rho)$ and $c_2(\rho)$, under various conditions on $\rho$, and with constructing mappings achieving optimal or near-optimal distortions.

In this paper we introduce and study certain subclasses of $l_1$-metrics. For each finite metric space $(X, \rho)$, we define a class $l_1^{\mathrm{dom}}(\rho)$ of metrics on$X$. Such metrics have been used implicitly before. They appear, e.g., in Bourgain's proof [4] showing that each $n$-point metric $\rho$ can be embedded into $l_2$ with distortion $O(\log n)$, and in Rao's [18] improvement of this construction for metrics of a special type (see Section 6 below for an example of such construction). The metrics in $l_1^{\mathrm{dom}}(\rho)$ have a relatively simple structure, yet many interesting properties, and a better understanding of such metrics may well prove rewarding.

Let $(X, \rho)$ be a finite metric space. A **$\rho$-dominated line metric** on $X$ is a metric $\tau$ induced by a nonexpanding (1-Lipschitz) embedding of $(X, \rho)$ into the real line. Explicitly, $\tau(x,y) = |\phi(x) - \phi(y)| \leq \rho(x,y)$ for some nonexpanding mapping $\phi\colon X \to \mathbb{R}$. The class $l_1^{\mathrm{dom}}(\rho)$ consists of all metrics on the ground set

$X$ that are convex combinations of $\rho$-dominated line metrics on $X$; that is,

$$l_1^{\mathrm{dom}}(\rho) = \left\{ \sum_{i=1}^{m} \alpha_i \tau_i \colon \alpha_i \geq 0, \sum \alpha_i = 1, \tau_1, \ldots, \tau_m \ \rho\text{-dominated line metrics} \right\}.$$

Clearly, each metric $\tau$ in $l_1^{\mathrm{dom}}(\rho)$ can also be viewed as induced by a nonexpanding embedding $\phi_\tau \colon (X, \rho) \to l_1^m$, whose $i$th coordinate $(\phi_\tau)_i$ is given by $\alpha_i$ times the nonexpanding embedding of $(X, \rho)$ into $\mathbb{R}$ inducing $\tau_i$. We define

$$c_1^{\mathrm{dom}}(\rho) = \inf\{D\colon \text{there is a } \tau \in l_1^{\mathrm{dom}}(\rho) \text{ such that } \phi_\tau \text{ is a } D\text{-embedding}\}.$$

In order to define $c_1^{\mathrm{dom}}(\rho)$ for an infinite metric space, we again take the supremum over all finite subspaces; also see the remark at the end of this section.

Two important properties of $c_1^{\mathrm{dom}}(\rho)$ were established by Bourgain [4]. First, his proof gives

$$c_1^{\mathrm{dom}}(\rho) \leq O(\log n)$$

where $n$, as usual, stands for the size of the underlying space. The bound is tight, since by [15], the shortest-path metric of a constant-degree expander graph achieves $c_1(\rho) = \Omega(\log n)$.*

The second property of $c_1^{\mathrm{dom}}(\rho)$ is

$$c_1^{\mathrm{dom}}(\rho) \geq c_2(\rho).$$

(Note that we have the *reverse* inequality for unrestricted embeddings into $l_1$: for all $\rho$, $c_1(\rho) \leq c_2(\rho)$; see, e.g., [15].) For completeness, here is an explanation. Let $\tau$ be the metric in $l_1^{\mathrm{dom}}(\rho)$ closest to $\rho$ (i.e., attaining $c_1^{\mathrm{dom}}(\rho)$), and assume that $\tau$ is a convex combination of $\rho$-dominated line metrics $\tau_i$:

$$\tau = \sum_{i=1}^{m} \alpha_i \tau_i.$$

Write $x_i = (\phi_\tau(x))_i$. Then the mapping

$$f\colon x \mapsto \left(\alpha_1^{1/2} x_1, \ldots, \alpha_m^{1/2} x_m\right) \in \mathbb{R}^m, \quad x \in X$$

from $(X, \rho)$ to $l_2^m$ has distortion at most $c_1^{\mathrm{dom}}(\rho)$. Indeed, on the one hand,

$$\|f(x) - f(y)\|_2^2 = \sum_{i=1}^{m} \alpha_i |x_i - y_i|^2 \leq \rho(x, y)^2,$$

* The notation $f = \Omega(g)$ is equivalent to $g = O(f)$, and $f = \Theta(g)$ means that both $f = O(g)$ and $f = \Omega(g)$.

while, on the other hand,

$$\|f(x)-f(y)\|_2=\left(\sum_{i=1}^{m}\alpha_i|x_i-y_i|^2\right)^{1/2}\geq\sum_{i=1}^{m}\alpha_i|x_i-y_i|\geq\frac{1}{c_1^{\mathrm{dom}}(\rho)}\rho(x,y).$$

We remark that a similar argument gives $c_1^{\mathrm{dom}}(\rho)\geq c_p(\rho)$ for all $p$, $1\leq p\leq\infty$.

The central issue of this paper is the estimation of the gap between $c_2(\rho)$ and $c_1^{\mathrm{dom}}(\rho)$. As was mentioned above, $c_2(\rho)\leq c_1^{\mathrm{dom}}(\rho)$. We show that for a finite metric space $(X,\rho)$ with $|X|=n$, we always have $c_1^{\mathrm{dom}}(\rho)=O(\sqrt{\log n})c_2(\rho)$, and that this bound is essentially tight. Using a probabilistic construction, we exhibit a family of metric spaces $\rho_n$ for which $c_1^{\mathrm{dom}}(\rho_n)=\Omega(\sqrt{\log n/\log\log n})c_2(\rho_n)$.

In order to obtain the lower bound, it is first shown that $c_1^{\mathrm{dom}}(S^n)=\Theta(\sqrt{n})$, where $S^n$ is the Euclidean $n$-sphere equipped with its geodesic metric. Then a discrete analogue of this result is obtained for a suitable finite random subset of $S^n$.

Finally, we also discuss two examples of non-obvious embeddings of special metric spaces into $l_1^{\mathrm{dom}}$.

Remarks on the definition of $c_1^{\mathrm{dom}}$ for infinite metric spaces. For an infinite metric space $(X,\rho)$, we have defined $c_1^{\mathrm{dom}}(\rho)$ as the supremum of $c_1^{\mathrm{dom}}(\tau)$ over finite submetrics $\tau$ of $\rho$; thus, we have postulated a "compactness" property. A natural way of introducing the corresponding set $l_1^{\mathrm{dom}}(\rho)$ of metrics seems to be the following. Let $\Lambda(\rho)$ denote the set of all $\rho$-dominated line metrics. Consider $\Lambda(\rho)$ as a subspace of the space $\mathbb{R}^{X\times X}$ with the product topology; in other words, with the topology of pointwise convergence of functions on $X\times X$. We define $l_1^{\mathrm{dom}}(\rho)$ as the closure of the convex hull of $\Lambda(\rho)$ in this topology (for $X$ finite, $\mathrm{conv}(\Lambda(\rho))$ is closed, and so the new definition agrees with the one given earlier for finite spaces).

Having defined $l_1^{\mathrm{dom}}(\rho)$ for an infinite $\rho$, we can now define $c_1^{\mathrm{dom}}(\rho)$ as the infimum of the distortions of the identity mapping $(X,\rho)\to(X,\tilde{\rho})$, where $\tilde{\rho}\in l_1^{\mathrm{dom}}(\rho)$. Let us check that this agrees with the previous definition of $c_1^{\mathrm{dom}}(\rho)$ via finite subspaces. Since the new definition gives no new metrics on finite subspaces, the distortion according to the new definition cannot be smaller than the one from the old definition. To see the opposite inequality, we first note that $l_1^{\mathrm{dom}}(\rho)$ is compact, being a closed subspace of the product of the closed intervals $[0,\rho(x,y)]$, $x,y\in X$. Suppose that $c_1^{\mathrm{dom}}(\tau)<D$ for all finite subspaces $(Y,\tau)$ of $(X,\rho)$. For each $(Y,\tau)$, define $F_Y$ as the set of all $\bar{\rho}\in l_1^{\mathrm{dom}}(\rho)$ that, when restricted to $Y$, distort $\tau$ by at most $D$. We observe that any $\tau$-dominated line metric on $Y$ can be extended to a $\rho$-dominated line metric on $X$; this follows

from McShane's theorem, stating that any 1-Lipschitz real function defined on a subspace of a metric space can be extended to a 1-Lipschitz function on the whole space (see, e.g., [20]). It follows that $F_Y \neq \emptyset$. Each $F_Y$ is closed, and for finite $Y_1, Y_2, \ldots, Y_k$, we have $F_{Y_1} \cap F_{Y_2} \cap \cdots \cap F_{Y_k} \supseteq F_{Y_1 \cup \cdots \cup Y_k} \neq \emptyset$. Hence the family $\{F_Y\}$ has a nonempty intersection, and any metric in the intersection distorts $\rho$ by at most $D$.

For some of the subsequent proofs, it is convenient to note that $l_1^{\mathrm{dom}}(\rho)$ as defined above contains "infinite convex combinations" of the $\rho$-dominated line metrics. Namely, if $\mu$ is a Borel probability measure supported on $\Lambda(\rho)$, then the metric $\tilde{\rho}$ given by $\tilde{\rho}(x,y) = \int_{\Lambda(\rho)} \tau(x,y)\,\mathrm{d}\mu(\tau)$ lies in $l_1^{\mathrm{dom}}(\rho)$.

Finally, we note that if $(X,\rho)$ is compact, then $l_1^{\mathrm{dom}}(\rho)$ coincides with the closure of $\mathrm{conv}(\Lambda(\rho))$ in the supremum norm on $C(X \times X)$ (uniform convergence). Indeed, the elements of $\mathrm{conv}(\Lambda(\rho))$, regarded as functions $X \times X \to \mathbb{R}$, are 1-Lipschitz with respect to the product metric on $X \times X$ derived from $\rho$ (the distance of pairs $(x,y)$ and $(z,t)$ is $\rho(x,z) + \rho(y,t)$). Thus, if $(X,\rho)$ is compact, pointwise convergence of these functions is the same as uniform convergence. We omit more detailed discussion of these "infinite-dimensional" aspects, since we are mainly interested in finite metrics.

## 2. The geodesic metric of $S^n$

We start with a simple observation about the $n$-dimensional Euclidean space. Let $S^n$ denote the $n$-dimensional unit sphere in $\mathbb{R}^{n+1}$, and let $\mu_n$ be the uniform probability measure on $S^n$.

CLAIM 2.1: $c_1^{\mathrm{dom}}(l_2^n) = O(\sqrt{n})$.

*Proof:* Let $\tau_\vartheta$ be the dominated line metric of $l_2^n$ obtained by the orthogonal projection of $\mathbb{R}^n$ on the line defined by the center of $S^n$ and a point $\vartheta \in S^n$. Consider the metric $\tilde{\rho}$ defined by

$$\tilde{\rho} = \int_{S^n} \tau_\vartheta \,\mathrm{d}\mu_n(\vartheta).$$

This is an element of $l_1^{\mathrm{dom}}(\rho)$, where $\rho$ is the Euclidean metric on $l_2^n$ (see the remarks at the end of Section 1). The distortion of the identity mapping with respect to the metrics $\rho$ and $\tilde{\rho}$ equals the expected length of projection of a random unit vector on a fixed line. It is well known (see, e.g., [17], Section 2) that this expectation is $\Theta(1/\sqrt{n})$, implying the claim. ∎

Combining Claim 2.1 with a simple observation that the Euclidean metric of $S^n$ shrinks its geodesic metric by at most $\pi/2$, one gets

COROLLARY 2.2: $c_1^{\mathrm{dom}}(S^n) = O(\sqrt{n})$.

Proving lower bounds is harder. Using the standard approach in such situations, let us first write $c_1^{\mathrm{dom}}(\rho)$ in a dual form, more suited for the task. For simplicity, we use the discrete notation. By duality of linear programming (Farkas' Lemma, or separation of disjoint convex sets by a hyperplane), we have

CLAIM 2.3: *For any finite metric space $(X, \rho)$,*

$$c_1^{\mathrm{dom}}(\rho) = \max_L \min_\tau \frac{L(\rho)}{L(\tau)},$$

*where the maximum is taken over all nonnegative linear forms $L(\delta) = \sum_{i,j\in X} a_{ij}\delta_{ij}$, and the minimum is taken over all $\rho$-dominated line metrics $\tau$.*

Thus, in order to obtain a lower bound on $c_1^{\mathrm{dom}}(\rho)$, it suffices to exhibit a single linear form $L$ such that $L(\rho)$ is sufficiently bigger than $\max_\tau L(\tau)$. The main difficulty of this approach lies in bounding the latter expression.

Our next goal is to obtain a lower bound on the distortion of the geodesic metric of $S^n$. We shall use the approach outlined above (in a version for an infinite metric space). In order to make it work, we use the classical Levy's Lemma about the metric structure of $S^n$. We follow [17], Chapter 2.

Given a set $K \subseteq S^n$ and $\varepsilon \geq 0$, define $K^\varepsilon$ as the set of all points $x \in S^n$ that are at (geodesic) distance at most $\varepsilon$ from $K$.

LEMMA 2.4 (Levy's Lemma): *Let $f\colon S^n \to \mathbb{R}$ be a continuous function, and denote by $M_f$ its median, i.e., a number such that both the sets $\{x\colon f(x) \leq M_f\}$ and $\{x\colon f(x) \geq M_f\}$ have measure $\geq \frac{1}{2}$. Let $A = \{x\colon f(x) = M_f\}$. Then*

$$\mu_n(A^\varepsilon) \geq 1 - \sqrt{\frac{\pi}{2}}e^{-\varepsilon^2(n-1)/2}.$$

This lemma is easily derived from two results which we recall below (as we will need them later); for more information see, e.g., [17].

THEOREM 2.5 (Isoperimetric inequality for sphere): *For any $\varepsilon \geq 0$ and any $K \subseteq S^n$ with $\mu_n(K) = \kappa$, $0 \leq \kappa \leq 1$,*

$$\mu_n(K^\varepsilon) \geq \mu_n(C_\kappa^\varepsilon),$$

*where $C_\kappa$ denotes a spherical cap of measure $\kappa$.*

LEMMA 2.6: *The measure of a spherical cap of geodesic radius $\pi/2+\varepsilon$ satisfies*

$$\mu_n(C^{\varepsilon}_{1/2}) \geq 1 - \sqrt{\frac{\pi}{8}}e^{-\varepsilon^2(n-1)/2}.$$

This is a "spherical analogue" of the well-known Chernoff inequality for the binomial distribution, and it can be proved by more or less straightforward calculations.

COROLLARY 2.7: *Let $f\colon S^n \to \mathbb{R}$ be a nonexpanding map. Then, for every $t \geq 0$, there exists an interval $I_t \subset \mathbb{R}$ of length $t$ such that*

$$\mu_n(f^{-1}I_t) \geq 1 - \sqrt{\frac{\pi}{2}}e^{-t^2(n-1)/8}.$$

The corollary follows from Lemma 2.4: let $I_t$ be the $t/2$-neighbourhood of $f(A) = M_f \in \mathbb{R}$ and note that $f(A^{t/2}) \subseteq I_t$ since $f$ is nonexpanding.

COROLLARY 2.8: *Let $f\colon S^n \to \mathbb{R}$ be a nonexpanding map. Then, for every $t \geq 0$,*

$$\mathrm{Prob}_{\mu_n\times\mu_n}\left[|f(x)-f(y)| \geq t\right] \leq \sqrt{2\pi}e^{-t^2(n-1)/8}.$$

The corollary follows from Corollary 2.7 by bounding the probability that two randomly and independently chosen points of $S^n$ are *not* both mapped to $I_t$ by $f$.

We can now prove a lower bound on the distortion of embedding $S^n$ into $l_1^{\mathrm{dom}}$, and show that Corollary 2.2 is tight:

THEOREM 2.9:

$$c_1^{\mathrm{dom}}(S^n) \geq \frac{\pi}{4}\sqrt{n-1}.$$

*Proof:* Following the general approach for establishing lower bounds via Farkas' Lemma, consider the following linear functional $L(\tau)$ over metrics $\tau$ on $S^n$:

$$L(\tau) = \int_{S^n}\int_{S^n} \tau(x,y)\,\mathrm{d}\mu_n\,\mathrm{d}\mu_n = \mathbf{E}_{\mu_n\times\mu_n}[\tau(x,y)].$$

The value of $L$ on the geodesic metric $\rho$ of $S^n$ is, clearly, $\pi/2$. Consider now an arbitrary nonexpanding map $f\colon S^n \to \mathbb{R}$ and the dominated line metric $\tau$ induced by $f$ on $S^n$. By a standard argument,

$$L(\tau) = \mathbf{E}[\tau(x,y)] = \int_0^\infty \mathrm{Prob}[\tau(x,y) \geq t]\,\mathrm{d}t.$$

Using Corollary 2.8 we get

$$L(\tau) \leq \int_0^\infty \sqrt{2\pi}e^{-t^2(n-1)/8}\,\mathrm{d}t = \sqrt{\frac{8\pi}{n-1}}\int_0^\infty e^{-y^2/2}\,\mathrm{d}y = \frac{2}{\sqrt{n-1}}.$$

If $\tilde{\rho}$ is a convex combination of $\rho$-dominated line metrics, the just derived inequality implies that $\tilde{\rho}$ distorts $\rho$ by at least $\sqrt{n-1}/2$ (we use the "easy direction" of Farkas' Lemma; it is obvious that if $\tilde{\rho}$ distorts $\rho$ by at most $D$, then $L(\tilde{\rho}) \geq L(\rho)/D$). Finally, any metric in $l_1^{\mathrm{dom}}(\rho)$ is a pointwise limit of such convex combinations $\tilde{\rho}$ (see the end of Section 1), and so its distortion is at least $\sqrt{n-1}/2$ as well. ∎

*Remark:* We have recently learned that Gromov [11], in his study of the concentration properties of metric spaces, uses a technique very similar to the one employed by us in getting the lower bounds. While Gromov focuses mainly on the *Observable Diameter* of metric spaces, our lower bounds are based on (using Gromov's terminology) the closely related *Observable Mean Distance.* Given a (compact) metric space $\rho$, the Observable Mean Distance of $\rho$ is defined as the maximum possible mean distance of a $\rho$-dominated line metric.

## 3. Finite metrics

This section is the heart of the paper; it deals with establishing sharp bounds on the gap between $c_2(\rho)$ and $c_1^{\mathrm{dom}}(\rho)$ as a function of the size of the underlying metric space.

We start with the simpler upper bound.

THEOREM 3.1: *Let $(X, \rho)$ be a finite metric space with $|X| = n$. Then $c_1^{\mathrm{dom}}(\rho) = O(\sqrt{\log n})c_2(\rho)$.*

*Proof:* Clearly, it suffices to show that if the metric space $(X, \rho)$ is $l_2$-embeddable, then $c_1^{\mathrm{dom}}(\rho) = O(\sqrt{\log n})$. Assume that this is the case. By the fundamental Johnson–Lindenstrauss Lemma [14], $(X, \rho)$ is also embeddable in $l_2^{\log n}$ with a constant distortion. Applying Claim 2.1 to $l_2^{\log n}$, we deduce the theorem. ∎

The lower bound is obtained via a discretization of Theorem 2.9. The line of reasoning will mimic that of the previous section.

Denote by $\mathrm{Cap}(a, r) \subseteq S^n$ a spherical cap of geodesic radius $r$ centered at $a$. The following lemma allows us to use the ideas from the continuous case in the discrete setting:

LEMMA 3.2: *Let $r$ be a positive number with $r < r_0$ for a suitable universal constant $r_0 > 0$. For any natural $n$, there exist $N = \Theta(1/r)^{n+1}$ and a covering of $S^n$ by $N$ caps of geodesic radius $r$ with the following properties:*

a. *The family of caps* $\{\mathrm{Cap}(a,r)\colon a \in X\}$ *covers* $S^n$.
b. *No point of* $S^n$ *is covered by more than* $\theta = O(n(\log n + \log \frac{1}{r}))$ *caps in this family.*
c. *The average geodesic distance between the points of* $X$ *is* $\pi/2 + o(1)$.

We shall delay the proof of the lemma until later, and show how it is used. Let $(X,\rho)$ be a subset of $S^n$ with properties as ensured by Lemma 3.2 for $r = 1/\sqrt{n}$; then $N = \Theta(n)^{0.5(n+1)}$ and $\theta = O(n\log n)$. The size of $X$ is $N$, and $\rho$ is induced by the geodesic metric of $S^n$. Consider the linear form

$$L(\delta) = \frac{1}{N^2}\sum_{i,j\in X}\delta_{ij} = \mathbf{E}[\delta(x,y)],$$

expressing the expected distance between two random points of $X$. By property c., $L(\rho) = \pi/2 + o(1)$. Our goal is to show that the value of $L(\tau)$ on any $\rho$-dominated line metric $\tau$ is at most $O\big(\sqrt{(\log n)/n}\big) = O\big(\sqrt{\log\log N/\log N}\big)$. This, of course, would imply the theorem.

Claim 3.3: *Let* $K$ *be a subset of* $X$. *Then*

$$\frac{1}{\theta}\cdot\frac{|K|}{N} \le \mu_n(K^r) \le \theta\cdot\frac{|K|}{N},$$

*where, as before,* $K^r$ *stands for the* $r$*-neighbourhood of* $K$.

*Proof:* Let $v_r$ denote the $\mu_n$-measure of a spherical cap of geodesic radius $r$. By the properties a. and b., we have $\theta \ge N\cdot v_r \ge 1$. Therefore,

$$\mu_n(K^r) \le |K|\cdot v_r \le \theta\cdot\frac{|K|}{N}.$$

On the other hand,

$$\mu_n(K^r) \ge \frac{1}{\theta}\cdot|K|\cdot v_r \ge \frac{1}{\theta}\cdot\frac{|K|}{N}. \qquad \blacksquare$$

Claim 3.4: *Let* $Q \subseteq X$ *be a subset of* $X$ *of size* $\ge N/2$, *and let* $P \subseteq X$ *be a set of points such that*

$$\rho(Q,P) \ge \varepsilon + 2r + \delta,$$

*where* $\delta = \sqrt{(2\ln 2\theta)/(n-1)}$. *Then* $|P|/N \le \theta\cdot e^{-\varepsilon^2(n-1)/2}$.

*Proof:* By Claim 3.3, we have

$$\mu_n(Q^r) \ge \frac{1}{\theta}\cdot\frac{|Q|}{N} \ge \frac{1}{2\theta}.$$

Using Lemma 2.6 on the complement of the cap $C_{1/2\theta}$, we calculate that the radius of $C_{1/2\theta}$ is at least $\pi/2-\delta$, and by the Isoperimetric Inequality 2.5, we obtain

$$\mu_n(Q^{r+\varepsilon+\delta}) \geq \mu_n(C_{1/2\theta}^{\varepsilon+\delta}) \geq v_{\pi/2+\varepsilon} \geq 1-\sqrt{\frac{\pi}{8}}e^{-\varepsilon^2(n-1)/2}.$$

Since $P^r$ and $Q^{r+\varepsilon+\delta}$ have disjoint interiors, we have

$$\mu_n(P^r) \leq 1-\mu_n(Q^{r+\varepsilon+\delta}) \leq \sqrt{\frac{\pi}{8}}e^{-\varepsilon^2(n-1)/2}.$$

The claim now follows from Claim 3.3. ∎

Finally, we can show a lower bound on the gap between $c_1^{\mathrm{dom}}(\rho)$ and $c_2(\rho)$ for our $(X,\rho)$.

THEOREM 3.5: $c_1^{\mathrm{dom}}(\rho) = \Omega(\sqrt{\log N/\log\log N})c_2(\rho)$.

*Proof:* As we have seen before, it suffices to show that the expected distance between two random points of any $\rho$-dominated metric line $\tau$ on $X$ is $\Omega(\sqrt{(\log n)/n})$. Recall that $N=\Theta(n)^{0.5(n+1)}$.

Arguing as in the derivation of Corollary 2.7 from Levy's Lemma (our Lemma 2.4), and keeping in mind $r=1/\sqrt{n}$, we conclude from Claim 3.4 that all but, say, an $n^{-2}$ fraction of points in $f(X)$ can be covered by an interval of length $B\sqrt{(\log n)/n}$ for a suitable constant $B$. Consequently, the probability that the $\tau$-distance between two randomly chosen points of $X$ exceeds $B\sqrt{(\log n)/n}$ is $O(n^{-2})$. Keeping in mind that the maximal $\tau$-distance between any two points of $X$ is at most $\pi$, we conclude that $\mathbf{E}_{X\times X}[\tau(x,y)]$ is at most

$$\begin{aligned} B\sqrt{(\log n)/n}+\pi\cdot\mathrm{Prob}\left[\tau(x,y)\geq B\sqrt{(\log n)/n}\right] &= B\sqrt{(\log n)/n}+O(n^{-2}) \\ &= O\big(\sqrt{(\log n)/n}\big). \end{aligned}$$

This concludes the proof of the theorem. ∎

## 4. Proof of Lemma 3.2

*Proof:* The proof is essentially the same as in Rogers [19] and Erdős and Rogers [8]. Since we have no explicit reference for the exact result we need here, we include a proof. The calculations are somewhat simpler than those in [19, 8], since a less precise bound is sufficient for our application.

In the sequel, we may suppose that $n$ is sufficiently large. Let $r$ be given, and set $\eta=r/n^2$. Let $v_x$ denote, as before, the $\mu_n$-measure of a spherical cap of geodesic radius $x$. Define $N$ as the smallest integer such that $N \cdot v_{r-\eta} \geq 14 \ln(n/v_\eta)$. Let $X = \{a_1, a_2, \ldots, a_N\}$ be a subset of $S^n$ obtained by choosing $N$ points independently and uniformly at random. We show that $X$ satisfies the requirements of the lemma with probability close to 1.

We start with a.

Let $B \subseteq S^n$ be the set of points not covered by any of $\text{Cap}(a_i, r-\eta)$, $i = 1, 2, \ldots, N$. The expected $\mu_n$-measure of $B$ equals the probability of a fixed point $x \in S^n$ being missed by all $\text{Cap}(a_i, r-\eta)$. By definition of $N$, the latter probability is

$$(1 - v_{r-\eta})^N < e^{-N \cdot v_{r-\eta}} \leq n^{-14} v_\eta.$$

Thus, by Markov's Inequality, the measure of $B$ is smaller than $v_\eta$ with probability at least $1 - n^{-14}$. In such case, the caps $\text{Cap}(a_i, r)$ cover the entire $S^n$, for if $x$ were not covered by any $\text{Cap}(a_i, r)$, we would have $\text{Cap}(x, \eta) \subseteq B$ implying $\mu_n(B) \geq v_\eta$.

We continue with b.

Set $\theta = \frac{5}{3} N \cdot v_{r+\eta}$. As we shall see later, this $\theta$ is indeed of the correct order of magnitude. The expected measure of the subset of points $S^n$ covered more than $\theta$-times by the caps $\text{Cap}(a_i, r+\eta)$, equals the probability of a fixed $x \in S^n$ being covered more than $\theta$-times. The number of $(r+\eta)$-caps covering $x$ behaves as a sum of $N$ independent 0/1 random variables, each of them attaining value 1 with probability $v_{r+\eta}$. By Chernoff's Bound (as in Alon and Spencer [1]; see the remark above Theorem A.12), the probability of more than $\theta$ caps covering $x$ is bounded above by

$$e^{-2N \cdot v_{r+\eta}/27} \leq e^{-2N \cdot v_{r-\eta}/27} \leq e^{-\ln(n/v_\eta)} = \frac{1}{n} v_\eta.$$

Thus, by Markov's Inequality, with probability at least $1 - 1/n$, the fraction of the points of $S^n$ that are covered more than $\theta$ times by the caps $\text{Cap}(a_i, r+\eta)$ is no more than $v_\eta$. Consequently, no point $x$ is covered more than $\theta$ times by the $r$ caps $\text{Cap}(a_i, r)$, since otherwise the entire $\text{Cap}(x, \eta)$ would be covered more than $\theta$ times by caps $\text{Cap}(a_i, r+\eta)$.

Finally, we show c.

Define $\Delta_{ij}$ as the distance between $a_i$ and $a_j$, $i < j$. Then each $\Delta_{ij}$ is a random variable with values in the interval $[0, \pi]$ and with distribution symmetric about $\pi/2$. Due to the uniformity of $S^n$, the $\Delta_{ij}$ are pairwise independent. Therefore

$$\mathbf{E}\Big[\sum_{i<j} \Delta_{ij}\Big] = \binom{N}{2} \pi/2 \quad \text{and} \quad \text{Var}\Big[\sum_{i<j} \Delta_{ij}\Big] = \sum_{i<j} \text{Var}[\Delta_{ij}] = O(N^2).$$

Using the Chebysheff Inequality we conclude that the average distance deviates from $\pi/2$ by more than $O(1/N)$ with probability $o(1)$.

It remains to show that $N$ and $\theta$ of **b.** are of the claimed order of magnitude. In [19], formula (7), it is proved that

$$v_r \geq \frac{1}{4\sqrt{n}} \sin^{n+1} r. \tag{1}$$

From the calculation preceding formula (5) in the same paper, it follows that for $r \leq s \leq r_0$ with a sufficiently small constant $r_0 > 0$,

$$\frac{v_s}{v_r} \leq 2\left(\frac{\sin s}{\sin r}\right)^{n+1}. \tag{2}$$

Recall that our $\eta$ is $r/n^2$. Using (2) and keeping in mind that $n$ is sufficiently large, we conclude that

$$\frac{v_{r+\eta}}{v_{r-\eta}} \leq 2\left(\frac{\sin(r+\eta)}{\sin(r-\eta)}\right)^{n+1} \leq 4.$$

Hence

$$\begin{aligned} N &\leq \frac{14\ln(n/v_\eta)}{v_{r-\eta}} \leq \frac{4}{v_r} \cdot 14\ln(n/v_\eta) \\ &= O(1)\cdot\sqrt{n}\cdot O\Big(\frac{1}{r}\Big)^{n+1} \cdot \ln\left(n^{1.5}\cdot O\Big(\frac{n^2}{r}\Big)^{n+1}\right) = O\Big(\frac{1}{r}\Big)^{n+1}, \end{aligned}$$

and

$$\begin{aligned} \theta &= \frac{5}{3}N\cdot v_{r+\eta} \leq \frac{5}{3}\cdot\frac{14\ln(n/v_\eta)}{v_{r-\eta}} v_{r+\eta} \\ &\leq \frac{5}{3}\cdot 56\cdot\ln\left(n^{1.5}\cdot O\Big(\frac{n^2}{r}\Big)^{n+1}\right) = O(n\log n + n\log(1/r)). \end{aligned}$$ ∎

## 5. The Euclidean cube

In the remaining two sections, we give two examples of non-obvious low-distortion embeddings of metric $\rho$ into $l_1^{\mathrm{dom}}(\rho)$ and a matching lower bound for one of them.

The first example is the $n$-cube $\{-1,1\}^n$ equipped with the Euclidean metric $\rho$. Somewhat surprisingly, it turns out that $c_1^{\mathrm{dom}}(\rho)$ is much smaller than $\sqrt{n}$.

THEOREM 5.1: *Consider the hypercube $\{-1,1\}^n$, and let $\rho$ be the Euclidean metric on it. Then*

$$c_1^{\mathrm{dom}}(\rho) = O(n^{1/4}).$$

*Proof:* Throughout this proof, by $\sqrt{t}$ we shall mean the function

$$\sqrt{t} = \mathrm{sign}(t)|t|^{1/2}, \quad t \in \mathbb{R}.$$

The approximating metric will be a convex combination of $2^n$ line metrics, each corresponding to some $\varepsilon \in \{-1,1\}^n$. The line metric corresponding to $\varepsilon$ is given by the embedding $\phi_\varepsilon \colon \{-1,1\}^n \to l_1^1$,

$$\phi_\varepsilon(x) = \sqrt{\langle x, \varepsilon\rangle},$$

where $\langle *, *\rangle$ denotes the scalar product. Let us start by showing that each of those metrics is $\rho$-dominated. Let $x, y \in \{-1,1\}^n$ be two vertices of the cube, and let $\Delta = y - x$. Then

$$|\phi_\varepsilon(x) - \phi_\varepsilon(y)| = \left|\sqrt{\langle x,\varepsilon\rangle} - \sqrt{\langle x,\varepsilon\rangle + \langle \Delta,\varepsilon\rangle}\right| \le |2\langle \Delta,\varepsilon\rangle|^{1/2} \le \|\Delta\|_2.$$

(The second inequality follows from

$$\left|\sqrt{t} - \sqrt{t+d}\right| \le \sqrt{|d|/2} - \sqrt{-|d|/2} = (2d)^{1/2}.$$

In the last inequality we have used the fact that each $\Delta_i$ can only take values in $\{-2,0,2\}$.)

Next, we show that for a constant fraction of the $\varepsilon$, we have

$$|\phi_\varepsilon(x) - \phi_\varepsilon(y)| \ge \Omega(n^{-1/4}) \cdot \|\Delta\|_2.$$

Indeed, it can be shown that for any two given vectors $x$ and $\Delta$ in $\mathbb{R}^n$, and for $\varepsilon \in \{-1,1\}^n$ chosen at random, the probability that both $|\langle x,\varepsilon\rangle| \le 2\|x\|_2$ and $|\langle \Delta,\varepsilon\rangle| \ge \frac{1}{2}\|\Delta\|_2$ is greater than $\frac{1}{4}$ (one needs only some rough estimations on the tails of binomial distibutions, obtainable by standard arguments; to avoid calculations see, e.g., [10]). Thus, for at least a quarter of the $\varepsilon$, we have

$$\begin{aligned}|\phi_\varepsilon(x) - \phi_\varepsilon(y)| &= \left|\sqrt{\langle x,\varepsilon\rangle} - \sqrt{\langle x,\varepsilon\rangle + \langle \Delta,\varepsilon\rangle}\right| \\ &\ge \left|\sqrt{2n^{1/2}} - \sqrt{2n^{1/2} + \frac{1}{2}\|\Delta\|_2}\right| = \Omega(n^{-1/4}) \cdot \|\Delta\|_2.\end{aligned}$$

This concludes the proof of the theorem. ∎

The bound of Theorem 5.1 is tight. A matching lower bound can be shown along the same lines as in the proof of Theorem 2.9. In order to avoid repetitions, we describe here the argument only in general lines.

Instead of working with $\{-1,1\}^n$, it will be more convenient to work with $\{0,1\}^n$. Let $\rho$ be the Euclidean metric of $\{0,1\}^n$. It is easy to see that the expected $\rho$-distance between two random points of $\{0,1\}^n$ is $\Theta(\sqrt{n})$. We shall argue that for any $\rho$-dominated metric $\tau$ on $\{0,1\}^n$, this expectation is $O(n^{1/4})$. By Claim 2.3, this would imply the desired lower bound.

Now, by the Isoperimetric Inequality for the Hamming cube, a discrete analogue of spherical Isoperimetric Inequality 2.5, the number of vertices of $\{0,1\}^n$ at Hamming distance $\leq \epsilon$ from a subset $A \subseteq \{0,1\}^n$ of size $\kappa$ is minimized when $A$ is a Hamming ball of size $\kappa$ (Harper [13]; also see, e.g., [3] for discussion).

Also, by the Chernoff Bound (see, e.g., the Appendix of [1]), which can be regarded as a discrete analogue of Lemma 2.6, the Hamming ball of radius $n/2+\epsilon$ contains at least $1 - e^{-2\epsilon^2/n}$ fraction of the vertices of $\{0,1\}^n$.

Combining these two facts and keeping in mind that the Euclidean metric of $\{0,1\}^n$ is precisely the square root of its Hamming metric, we conclude that for any $A \subset \{0,1\}^n$ containing half or more of the vertices, and any $\delta > 0$, the fraction of vertices whose $\rho$-distance from $A$ does not exceed $\delta$ is at least $1-e^{-2\delta^4/n}$. This, in turn, allows us to obtain the analogues of Levy's Lemma 2.4 and of Corollary 2.7 for a nonexpanding function $f\colon \{0,1\}^n \to \mathbb{R}$, and conclude that the inverse image of the $t/2$-neighbourhood of $f$'s median contains at least $1 - 2e^{-\delta^4/16n}$ fraction of the vertices of $\{0,1\}^n$. This implies an analogue of Corollary 2.8, namely that for any $\rho$-dominated line metric $\tau$ and any $t > 0$, the probability that the $\tau$-distance between two random vertices of $\{0,1\}^n$ exceeds $t$ is at most $4e^{-\delta^4/16n}$. We conclude that the expectation of this distance is at most

$$\int_0^\infty 4e^{-t^4/16n}\,\mathrm{d}t = O(n^{1/4}). \qquad \blacksquare$$

*Remark:* Compare with Gromov [11], pp. 195–196, where a similar lower bound is shown for the Observable Diameter of the Euclidean cube.

## 6. The complete binary tree

Our second example is the complete binary tree $T_n$ of height $n$. Let $\rho$ be the shortest-path metric of $T_n$. Following the ideas of [18] (a modification of Bourgain's sampling method), we show that $c_1^{\mathrm{dom}}(\rho) = O(\log n)$ (note that we are dealing with a space with $2^n$ vertices, so the distortion is $\log\log$ of the number of vertices). This construction illustrates nicely how the sampling works. Let us remark that $c_2(\rho) = \Theta((\log n)^{1/2})$ [5] (also see [16]). Currently, we do not know what is the correct order of magnitude of $l_1^{\mathrm{dom}}(\rho)$.

Consider $T_n$ as a continuous structure, in which every edge is a segment of length 1. Assume for simplicity that $n$ is a power of 2. Let $r$ be the root of $T_n$. Let $i$ take values $2^0, 2^1, 2^2, \ldots$ up until $n$. For each $i$ and each $\alpha \in [0, i)$ define $X_{i,\alpha}$ as the set of all points $x$ in the continuous $T_n$, such that $\rho(x, r) \equiv \alpha (\bmod i)$. For every $X_{i,\alpha}$, consider all the connected components of $T_n \setminus X_{i,\alpha}$, and for a vertex $v \in V(T_n)$, let $C_v$ be the component containing $v$. Let $\varepsilon$ be a vector of $\pm 1$s assigning a sign to each component of $T_n \setminus X_{i,\alpha}$. For each such $\varepsilon$ (and for each $i$ and $\alpha$), we define a mapping $\phi = \phi_{i,\alpha,\varepsilon}\colon V(T_n) \to \mathbb{R}$ by

$$\phi(v) = \varepsilon(C_v) \cdot \rho(v, X_{i,\alpha}).$$

Let $\tau_{i,\alpha,\varepsilon}$ be the corresponding line metric, and let $\tau$ be the average over all such metrics. That is, every $i$ gets the same weight $\approx 1/\log_2 n$; for a fixed $i$, $\alpha$ is uniformly distributed; and for fixed $i$ and $\alpha$, all possible sign assignments are chosen with equal probability. We claim that $\tau$ has the desired properties.

Indeed, the triangle inequality for $\rho$ implies that each $\tau_{i,\alpha,\varepsilon}$ is a $\rho$-dominated line metric. On the other hand, let $x, y \in T_n$ be a pair of points at distance $d$, and let $j$ be the largest power of 2 such that $j \le d$. Thus, $2j > d \ge j$. We claim that the average distance between $x$ and $y$ taken over all metrics of the type $\tau_{j,\alpha,\varepsilon}$ is $\Theta(d)$. This claim implies that $\tau$ distorts $\rho$ by at most $\log n$.

Observe that $x$ and $y$ are separated by each $X_{j,\alpha}$. If it happens that, on the one hand, $\rho(x, X_{j,\alpha}) \ge j/4$ and, on the other hand, $x$ and $y$ get assigned different signs by $\varepsilon$, we obtain $\tau_{j,\alpha,\varepsilon}(x, y) \ge j/4 \ge d/8$. However, the two events are independent and each happens with probability 1/2. Hence, averaging over all the $\alpha$ and the $\varepsilon$, we get a contribution of at least $\frac{1}{32}d$ for $i = j$. ∎

*Remark:* In the introduction, we have indicated how an embedding of a metric $\rho$ into $l_1^{\mathrm{dom}}(\rho)$ with distortion $D$ yields an embedding of $\rho$ into $l_2$ with distortion at most$D$. The embedding into $l_1^{\mathrm{dom}}(\rho)$ with $O(\log n)$ distortion constructed in the proof just given is of a special type, and it actully yields an embedding into $l_2$ with distortion $O(\sqrt{\log n})$ only (thus, we get another proof, different from [5], [16], of the tight upper bound for embedding the complete binary tree into $l_2$). The reason is that, for each pair $x, y$ of points, the contribution of the coordinates in the embedding is "unevenly distributed"—there are about $1/\log n$ fraction of the dominated line metrics, each of them placing $x$ and $y$ at distance $\Omega(\rho(x, y))$. Such a phenomenon was used by Rao [18] for low-distortion embeddings of certain graphs (most notably, of planar graphs).

## 7. Conclusion

The main driving force behind this paper is the hope that bounding $c_1^{\mathrm{dom}}(\rho)$ might be easier than bounding $c_2(\rho)$, and thus good understanding of the structure of $l_1^{\mathrm{dom}}(\rho)$ should prove helpful for understanding embeddings into $l_2$. The properties of the polytope $l_1^{\mathrm{dom}}(\rho)$ deserve investigation. We have hardly touched upon the computational aspects of the theory. It is natural to ask whether there exists an efficient procedure for computing $l_1^{\mathrm{dom}}(\rho)$, or even for deciding whether a given $\tau$ belongs to $l_1^{\mathrm{dom}}(\rho)$. All these questions await future study.

### References


[1] N. Alon and J. Spencer, *The Probabilistic Method,* Wiley, New York, 1992.

[2] Y. Bartal, *On approximating arbitrary metrics by tree metrics,* in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing,* ACM Press New York, 1998, pp. 161–168.

[3] B. Bollobás, *Martingales, isoperimetric inequalities and random graphs,* in *52. Combinatorics, Eger (Hungary), Colloquia Mathematica Societatis János Bolyai,* North-Holland, Amsterdam, 1987, pp. 113–139.

[4] J. Bourgain, *On Lipschitz embedding of finite metric spaces in Hilbert space,* Israel Journal of Mathematics **52** (1985), 46–52.

[5] J. Bourgain, *The metrical interpretation of superreflexivity in Banach spaces,* Israel Journal of Mathematics **56** (1986), 222–230.

[6] P. Enflo, *On a problem of Smirnov,* Arkiv för Matematik **8** (1969), 107–109.

[7] P. Enflo, *On the nonexistence of uniform homeomorphisms between $L_p$-spaces,* Arkiv för Matematik **8** (1969), 103–105.

[8] P. Erdős and C. A. Rogers, *Covering space with convex bodies,* Acta Arithmetica **7** (1962), 281–285.

[9] U. Feige, *Approximating the bandwidth via volume respecting embeddings,* in *Proceedings of the 30th Annual ACM Symposium on Theory of Computing,* ACM Press New York, 1998, pp 90–99.

[10] D. Gillman, R. Permantel and Y. Rabinovich, *An inverse Chernoff bound,* manuscript, Haifa University, 1999; submitted to Information Processing Letters.

[11] M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces,* Birkhäuser, Boston, 1999.

[12] A. Gupta, I. Newman, A. Sinclair and Y. Rabinovich, *Cuts, trees and $l_1$ embeddings of graphs,* in *Proceedings of the 40th IEEE Symposium on Foundations of Computer Science,* IEEE Computer Society Press, Los Alamitos, CA, 1999, pp. 399–408.

[13] L. H. Harper, *Optimal numberings and isoperimetric problems on graphs,* Journal of Combinatorial Theory **1** (1966), 385–393.

[14] W. B. Johnson and J. Lindenstrauss, *Extentions of Lipschitz mappings into a Hilbert space,* Contemporary Mathematics **26** (1984), 189–206.

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

[16] J. Matoušek, *On embedding trees into uniformly convex Banach spaces,* Israel Journal of Mathematics **114** (1999), 221–237.

[17] V. Milman and G. Schechtman, *Asymptomic Theory of Finite Dimensional Spaces,* Lecture Notes in Mathematics **1200**, Springer-Verlag, Berlin, 1986.

[18] S. B. Rao, *Small distortion and volume preserving embeddings for planar and Euclidean metrics,* in *Proceedings of the 15th Annual ACM Symposium on Computational Geometry*, ACM Press, New York, NY, 1999, pp. 300–306.

[19] C. A. Rogers, *Covering a sphere with spheres,* Mathematika **10** (1963), 157–164.

[20] J. H. Wells and L. R. Williams, *Embeddings and Extensions in Analysis,* Springer, Berlin, 1975.