EMBEDDINGS OF LEVY FAMILIES INTO BANACH SPACES

M. RUDELSON AND R.VERSHYNIN

Abstract. We prove that if a metric probability space with a usual concentration property embeds into a Banach space X, then X has a proportional Euclidean subspace. In particular, this yields a new characterization of weak cotype 2. We also find optimal lower estimates on embeddings spaces with concentration properties (i.e. uniformly convex spaces) into l_∞^k, thus providing an "isomorphic" extension to results of Gromov-Milman and also generalizing estimates of Carl-Pajor and Gluskin.

1. Introduction

The concentration of measure phenomenon in various classes of probability metric spaces is a remarkable and well known theme in Geometric Functional Analysis. Discovered by V. Milman, it has been crucial in proofs of many results in the asympthotic theory of finite dimensional normed spaces.

Consider a probability metric space (T, μ, d) with the following concentration property for some constant c. For every $\varepsilon > 0$ and every subset A of T of measure at least $1/2$, the ε-inflation $A_\varepsilon = \{t \in T : d(t, A) < \varepsilon\}$ has measure at least $1 - 4\exp(-c\varepsilon^2 n)$. Here n is a parameter, usually an integer. If n varies, $n = 1, 2, \ldots$, then the family of such metric probability spaces (T_n, μ_n, d_n) is called a *Lèvy family*. The constant 4 is not important in this definition.

The notion of Lèvy family, introduced by M. Gromov and V. Milman [Gr-M], is by now standard, as it covers many natural families of spaces. Important examples of Lèvy families include the euclidean spheres $(S^{n-1}, \sigma_n, \rho_n)$ with the normalized geodesic distance and the normalized Lebesgue measure, the orthogonal groups $(O(n), \mu_n, \rho_n)$ with the Hilbert-Schmidt metric and the normalized Haar measure (and all homogeneous spaces of $O(n)$, like Stiefel manifolds and Grassmanian manifolds). A remarkable class of discrete Lèvy families is given by the powers $(T^n, \mu^{\otimes n}, d_n)$, where (T, μ) is arbitrary probability space, $\mu^{\otimes n}$

The research of the second author was supported in part by NSF grants DMS-9896342 and DMS-0070458.

is the product measure, and d_n is the normalized Hamming distance. More examples and references can be found in [M-S] and [Ta].

Let (T_n, μ_n, d_n) be a Lèvy family. To eliminate a trivial case, where the whole measure is concentrated in one atom or in its small neigborhood, let us assume that the ε-neighborhood of any point in T_n has measure smaller than $1-\delta$, for some positive ε and δ independent of n. We prove that if (T_n, d_n) can be C-Lipschitz embedded into an n-dimensional Banach space X, then X has a Euclidean subspace of dimension proportional to n. In other words, the Euclidean sphere S^{k-1}, which itself is a member of a Levy family, Lipschitz embeds into X (with $k \sim \dim X$). This result highlights the importance of the concentration of measure phenomenon in the Euclidean sphere: if *some* metric space with a standard concentration property embeds into X, then so does the euclidean sphere.

This result gives a new characterization for the Banach spaces of weak cotype 2. Recall that X has weak cotype 2 iff there exist constants c_1, c_2 such that every finite dimensional subspace Y of X contains in turn a subspace c_1-isomorphic to l_2^n with $n > c_2 \dim Y$. In particular, S^{n-1} Lipschitz embeds into Y with n proportional to $\dim Y$. The result stated above yields that if this definintion of weak cotype 2 holds for *some* Levy family (T_n, μ_n, d_n) instead of the Euclidean spheres S^{n-1}, then it must also hold for the Euclidean spheres, i.e. X must have weak cotype 2.

Our second result states that Levy families poorly embed into l_∞^k. If (T_n, μ_n, d_n) is a regular Lèvy family, then for any map $F : T_n \to l_\infty^k$

$$\|F\|_{Lip} \|F^{-1}|_{F(T_n)}\|_{Lip} \geq c\sqrt{\frac{n}{\log(2+k/n)}}. \tag{1}$$

This bound is optimal, which can be seen from the recent "isomorphic Dvoretzky theorem" [M-S 98]: every k-dimensional Banach space contains an n-dimensional subspace $\psi(k,n)$-isomorphic to the Euclidean space, where $\psi(k,n)$ denotes the right side of (1). If $T = S^{n-1}$ is a Euclidean sphere and F is a linear operator, the estimate (1) was proved independently by J. Bourgain, J. lindenstrauss and V. Milman [B-L-M], B. Carl, A. Pajor [C-P] and E. Gluskin [G 89]: any n-dimensional subspace of l_∞^k has distance to l_2^n at least $\psi(k,n)$.

Our approach to (1) easily carries over to probability metric spaces with different concentration behavior. An important example is given by the class of uniformly convex Banach spaces. M. Gromov and V. Milman [Gr-M] (see [Schm]) proved that the sphere S_X of an n-dimensional uniformly convex space X has the same concentration property as a Levy family, but with the exponent p instead of 2 (p depends in a natural way on the degree of the uniform convexity).

Then an analogue of (1) for the metric space S_X proves that every n-dimensional subspace of l_∞^k has distance to X at least $c(n/\log(2+k/n))^{1/p}$. In the range $1 \le n \le c \log k$ this estimate is known from the paper of M. Gromov and V. Milman [Gr-M], so our result can be viewed as the "isomorphic" extension of their theorem.

It is worthwhile to note that in this paper we make no restrictions on the nature of the metric space besides the measure concentration property. This differs our perspective from the earlier results on embedding finite metric spaces into normed spaces, where different *particular* classes of metric spaces were considered: a generic n point metric space [Bo], [Ma 95], [J-L-S], [Ma 97], certain classes of graphs (expanders, trees) [L-L-R], [Ma 97], [Ma 99].

2. Norms of Subgaussian Random Vectors

Let $T = (T, \mu, d)$ be a probability metric space and A be a subset of T. For $\varepsilon > 0$, the ε-neighborhood of A is defined as $A_\varepsilon = \{x \in X : \rho(x, A) \le \varepsilon\}$. We define the concentration function of X as

$$\alpha(T, \varepsilon) = 1 - \inf\{\mu(A_\varepsilon) : A \subset T \text{ with } \mu(A) \ge 1/2\}. \tag{2}$$

Then a family $(T_n, \mu_n, d_n)_{n=1}^\infty$ of probability metric spaces is a *Lèvy family* with constant $c > 0$ if

$$\alpha(T_n, \varepsilon) \le 4 \exp(-c\varepsilon^2 n) \quad \text{for all } \varepsilon > 0 \tag{3}$$

and for all $n = 1, 2, \dots$

A natural nondegeneracy condition, which we will assume in some results, is the following: there exists positive constants ε and δ such that ε-neighborhood of any point in T_n has measure smaller than $1 - \delta$, $n = 1, 2, \dots$ Such Levy families are called *regular.*

It is a standard fact that if (T_n, μ_n, d_n) is a Lèvy family then every 1-Lipschitz function $F : T_n \to \mathbf{R}$ concentrates, i.e.

$$\mu_n\{|F - \mathbb{E}F| > \varepsilon\} \le 8 \exp(-c\varepsilon^2 n) \qquad \text{for all } \varepsilon > 0 \tag{4}$$

(see [M-S]).

This concentration property is convenient to be described in terms of the norm in the space L_{ψ_2}. The norm of a function f in $L_{\psi_2}(T, \mu)$ is the minimal number λ such that $\mathbb{E}\exp(f^2/\lambda^2) \le 4$. Equivalently, it is the minimal number λ such that

$$\mu\{|f| > s\} \le 8 \exp(-\lambda^2 s^2) \quad \text{for all } s > 0.$$

We see from (4) that if (T_n, μ_n, d_n) is a Lèvy family then for any mean zero 1-Lipschitz function $f : T_n \to \mathbf{R}$

$$\|f\|_{L_{\psi_2}(T_n,\mu_n)} \le c_1 n^{-1/2}, \tag{5}$$

(where $c_1 = c^{-1/2}$).

This easily generalizes to Banach space valued functions. Let X be a Banach space, and $F : T_n \to X$ be a 1-Lipschitz mean zero map. Then for every $x^* \in B_{X^*}$ the function $f : T_n \to \mathbf{R}$ defined as $f(\omega) = \langle F(\omega), x^* \rangle$, $w \in T_n$, is a 1-Lipschitz scalar valued mean zero function. Therefore by (4) $\|\langle F, x^* \rangle\|_{L_{\psi_2}(T_n,\mu_n)} \le c_1 n^{-1/2}$ for every $x^* \in B_{X^*}$.

This motivates the following definition. Let (T, μ, d) be a probability metric space and X be a Banach space. We will call a map $F : T \to X$ a *subgaussian random vector* if

$$\|\langle F, x^* \rangle\|_{L_{\psi_2}(T,\mu)} \le 1 \quad \text{for every } x^* \in B_{X^*}. \tag{6}$$

A "canonical" example of a subgaussian random vector is the standard Gaussian vector g in $(\mathbf{R}^n, \|\cdot\|_X)$, whose coordinates are independent $N(0,1)$ random variables, and the unit Euclidean ball B_2^n is contained in B_X. In this case, up to an absolute constant, $\frac{1}{\sqrt{n}}\mathbb{E}\|g\| \sim M_X$, where M_X is known as the M-estimate of X, $M_X = \int_{S^{n-1}} \|x\|_X \, d\sigma_n(x)$. The considered example is in some sense extremal. The main result in this section is

Theorem 1. *Let $X = (\mathbf{R}^n, \|\cdot\|_X)$ be a Banach space such that the maximal volume ellipsoid in B_X is the standard Euclidean ball B_2^n. Let F be a subgaussian random vector in X, i.e. (6) holds. Then*

$$\frac{1}{\sqrt{n}}\mathbb{E}\|F\| \le C M_X^{1/2}.$$

Here, as well as in the rest of the paper, C denotes absolute constants (possibly different in different places).

Proof. We write the average of $\|F\|_X$ as the expectation of the supremum of a sub-gaussian process:

$$\mathbb{E}\|F\|_X = \mathbb{E} \sup_{x^* \in B_{X^*}} \langle F, x^* \rangle.$$

We will cover B_{X^*} by translates of small Euclidean balls. Fix an $\varepsilon > 0$. By Sudakov's inequality ([Le-Ta] 3.3),

$$\Big(\log N(B_{X^*}, B_2^n, \varepsilon) \Big)^{1/2} \le C\varepsilon^{-1} \mathbb{E}\|g\|_X$$

($N(B_{X^*}, B_2^n, \varepsilon)$ denotes the minimal number of translates of εB_2^n needed to cover B_{X^*}). Let $\mathcal{N}$ be the set of points guaranteed by this entropy bound, i.e.

$$B_{X^*} = \mathcal{N} + \varepsilon B_2^n \quad \text{and} \quad (\log|\mathcal{N}|)^{1/2} \leq c\varepsilon^{-1}\mathbb{E}\|g\|_X.$$

Then

$$\mathbb{E} \sup_{x^* \in B_{X^*}} \langle F, x^* \rangle = \mathbb{E} \sup_{x^* \in \mathcal{N}} \langle F, x^* \rangle + \varepsilon \cdot \mathbb{E} \sup_{x^* \in B_2^n} \langle F, x^* \rangle. \tag{7}$$

We can certainly assume that $\mathcal{N} \subset B_{X^*}$, so by (6) we have

$$\|\langle F, x^* \rangle\|_{L_{\psi_2}(T,\mu)} \leq 1 \quad \text{for all } x^* \in \mathcal{N}. \tag{8}$$

It is easy to see that (8) alone implies that $\mathbb{E}\sup_{x^*\in\mathcal{N}}\langle F, x^*\rangle \leq C(\log|\mathcal{N}|)^{1/2}$ (see [Le-Ta] 3.1). Thus the first summand in (7) is bounded by

$$C\varepsilon^{-1}\mathbb{E}\|g\|_X.$$

Next, the second summand in (7) is

$$\mathbb{E} \sup_{x^* \in B_2^n} \langle F, x^* \rangle = \mathbb{E}\|F\|_2$$

(by $\|\cdot\|_2$ we denote the euclidean norm). We use the John's decomposition of the identity on X. Namely, if B_2^n is the maximal volume ellipsoid inscribed in B_X, then the identity operator on X can be decomposed as

$$id_X = \sum_{j=1}^m x_j \otimes x_j,$$

where $x_j/\|x_j\|_X$ are contact points, i.e. $\|x_j\|_X = \|x_j\|_{X^*} = \|x_j\|_2$, $j = 1, \ldots, m$, and $\sum_{j=1}^m \|x_j\|_2^2 = n$. We have

$$\begin{aligned}
\mathbb{E}\|F\|_2 &\leq (\mathbb{E}\|F\|_2^2)^{1/2} = \Big(\mathbb{E}\sum_{j=1}^m \langle F, x_j\rangle^2\Big)^{1/2} \\
&= \Big(\sum_{j=1}^m \|\langle F, x_j\rangle\|_{L_2(T,\mu)}^2\Big)^{1/2} \\
&\leq C\Big(\sum_{j=1}^m \|\langle F, x_j\rangle\|_{L_{\psi_2}(T,\mu)}^2\Big)^{1/2} \\
&\leq C\Big(\sum_{j=1}^m \|x_j\|_{X^*}^2\Big)^{1/2} \qquad \text{by (8)} \\
&= C\Big(\sum_{j=1}^m \|x_j\|_2^2\Big)^{1/2} = Cn^{1/2}.
\end{aligned}$$

As a consequence, (7) is bounded by

$$C(\varepsilon^{-1}\mathbb{E}\|g\|_X + \varepsilon n^{1/2}).$$

Therefore

$$\frac{1}{\sqrt{n}}\mathbb{E}\|F\|_X \le C\Big(\varepsilon^{-1}\frac{\mathbb{E}\|g\|_X}{\sqrt{n}} + \varepsilon\Big) \le C(\varepsilon^{-1}M_X + \varepsilon).$$

Now taking $\varepsilon = M_X^{1/2}$ we obtain the required estimate. The proof is complete.

A consequence for Levy families follows immediately.

Corollary 2. *Let $X = (\mathbf{R}^n, \|\cdot\|_X)$ be a Banach space such that B_2^n is the maximal volume ellipsoid in B_X. Let (T, d, μ) be the n-th term of a Levy family, i.e.* (3) *is satisfied. Then for any* 1*-Lipschitz mean zero map $F : T \to X$*

$$\mathbb{E}\|F\|_X \le C_1 M_X^{1/2},$$

where C_1 depends only on the constant c in the definition of Levy family.

It is a difficult question whether the estimates obtained are optimal. In particular, it is natural to ask whether the bound $CM_X^{1/2}$ can be improved to CM_X. One of the ways to improve it is to replace the John ellipsoid by an ellipsoid $\mathcal{E} \subset B_X$ for which M_X is maximal. An example in the paper by A. A. Giannopoulos, V. D. Milman and M. Rudelson [G-M-R] shows that this ellipsoid can be very far from the John's one.

Corollary 2 easily modifies to a dimension of X different from n.

Corollary 3. *Let $X = (\mathbf{R}^m, \|\cdot\|_X)$ be a Banach space, and all the remaining assumptions of Corollary 2 hold. Then*

$$\mathbb{E}\|F\|_X \le C_1 \left(\frac{m}{n}\right)^{1/2} M_X^{1/2}.$$

Proof. We modify the metric on T by setting $d'(t,s) = (n/m)^{1/2} d(t,s)$ for $t, s \in T$. Then (T, μ, d') is an m-th term of a Lèvy family, and F is an $(m/n)^{1/2}$-Lipschitz map from (T, μ, d') to X. Applying Corollary 2 to $(n/m)^{1/2}F$ we complete the proof.

Corolary 2 naturally applies to Lipschitz embeddings. As usual, a one-to-one map $F : (T, d) \to X$ is called an M-Lipschitz embedding if $\|F\|_{Lip}\|F^{-1}|_{F(T)}\|_{Lip} \le M$. We will be interested in estimating $k(X)$, the maximal dimension k of a subspace of X that is 2-isomorphic to l_2^k.

Corollary 4. *Let X be an n-dimensional Banach space, and (T, μ, d) be an n-th term of a regular Lèvy family. Assume that (T, d) can be M-Lipschitz embedded into X. Then*

$$k(X) \ge C(c, \varepsilon, \delta, M)n,$$

where c, ε, δ are the constants in the definition of the regular Lèvy family.

Proof. We choose the Euclidean structure on X as in Corollary 2. By [M-S] 4.2,

$$k(X) \geq CM_X^2 n, \tag{9}$$

and Corollary 2 will provide a lower bound for M_X.

Indeed, consider the map $\Phi = (F - \mathbb{E}F)/\|F\|_{Lip}$. Then Φ is mean zero, 1-Lipschitz, and $\|\Phi^{-1}|_{\Phi(T)}\|_{Lip} \leq M$. Let Φ' be an independent copy of Φ; then

$$\begin{aligned} \mathbb{E}\|\Phi\|_X &\geq \mathbb{E}\|\Phi - \Phi'\|_X \geq \frac{1}{M} \int_{T \times T} d(\omega, \omega') \, d\mu(\omega) \, d\mu(\omega') \\ &\geq \frac{1}{M} \varepsilon\delta. \end{aligned}$$

Then by Corollary 2 we have $M_X^{1/2} \geq C_1^{-1} \frac{1}{M} \varepsilon\delta$ which, when combined with (9), completes the proof.

3. Weak Cotype 2

One of several equivalent definitions of weak cotype 2 of a Banach space X is through a saturation of X by finite dimensional euclidean subspaces in the following sense. There are constants $\alpha, M > 0$ such that for every n and every subspace Y of X there exist a further subspace $Y \subset X$ with $k = \dim Y \geq \alpha n$, which is M-isomorphic to l_2^k (see [Pi 89]).

Our aim is to show that (the sphere of) the space l_2^k in this definition can be replaced by the k-th term of any regular Lèvy family. The linear embedding is replaced naturally by a Lipschitz embedding or, more generally, by a semi-Lipschitz embedding.

Definition 5. *Let (T_n, d_n) be a sequence of metric spaces, and X be a Banach space. Suppose we have for each n a one-to-one map $F_n : T_n \to X$. We call the family (F_n) a* semi-Lipschitz embedding *if*

(i) $\sup_n \|F_n\|_{Lip} < \infty$*;*

(ii) The family of maps (F_n^{-1}) defined on the images of F_n is equicontinuous.

We say that a family of metric spaces (T_n, d_n) *semi-Lipschitz saturates* a Banach space X if there is a constant $\alpha > 0$ such that for every sequence of subspaces (X_n) of X with $\dim X_n \geq \alpha n$ there is a semi-Lipschitz embedding (F_n) of (T_n, d_n) into X so that $F_n(T_n) \subset X_n$ for all n.

Theorem 6. *Suppose X is a Banach space, and there exists a regular Levy family which semi-Lipschitz saturates X. Then X has weak cotype 2.*

Proof. Consider a regular Levy family (T_n, d_n, μ_n) with regularity constants $\varepsilon, \delta > 0$, which saturates X. Let X_n be any subspaces of X with $\dim X_n \geq \alpha n$, and consider the corresponding semi-Lipschitz embedding $(F_n : T_n \to X_n)$. Since (F_n^{-1}) is equicontinuous, there exists a $\gamma > 0$ such that

$$\|F_n(\omega) - F_n(\omega')\|_{X_n} \geq \gamma \quad \text{whenever } d_n(\omega, \omega') \geq \delta.$$

Let F_n' be an independent copy of F_n; then

$$\begin{aligned} \mathbb{E}\|F_n - F_n'\|_{X_n} &\geq \gamma \cdot \mu_n \times \mu_n\{\|F_n(\omega) - F_n(\omega')\|_{X_n} \geq \gamma\} \\ &\geq \gamma \cdot \mu_n \times \mu_n\{d_n(\omega, \omega') \geq \delta\} \\ &\geq \gamma\varepsilon. \end{aligned}$$

Then, as in the proof of Corollary 4, $k(X_n) \geq C(c, \varepsilon, \delta, \gamma)n$. Thus X has weak cotype 2.

In general, it is impossible to interchange Lipshitzness and equicontinuity properties of F_n and F_n^{-1} in the definition of the semi-Lipschitz embedding. This is illustrated by the next two propositions.

Consider the discrete cube $C_2^n = \{-1, 1\}^n$, endowed with the normalized Hamming metric $d_n(x, y) = \frac{1}{2n}|\{i : x(i) \neq y(i)\}|$.

Proposition 7. *C_2^n cannot be semi-Lipschitz embedded into any normed space with type $p > 1$.*

Proof. This follows from a result of J. Bourgain, V. Milman and H. Wolfson [B-M-W]. Assume there is such an embedding $F : C_2^n \to X$, and put $X_\varepsilon = F(\varepsilon)$ for all $\varepsilon \in C_2^n$. We will show that X fails to have metric type $p > 1$. Since $d_n(\varepsilon, -\varepsilon) = 1$, it follows from the equicontinuity of the family (F_n^{-1}) that $\|X_\varepsilon - X_{-\varepsilon}\|_X \geq \delta$, where δ is some positive constant independent of n. Given a vertex $\varepsilon \in C_2^n$, let $\varepsilon[i]$ be the vertex in C_2^n differing from ε in the i-th coordinate only. The unordered pair $(\varepsilon, \varepsilon[i])$ is called an *edge*. There are $n2^{n-1}$ edges in C_2^n. We have $d_n(\varepsilon, \varepsilon[i]) = 1/n$, hence $\|X_\varepsilon - X_{\varepsilon[i]}\|_X \leq 1/n$. Then

$$D := \Big(\sum_{\varepsilon} \|X_\varepsilon - X_{-\varepsilon}\|_X^2 \Big)^{1/2} \geq 2^{n/2}\delta$$

and

$$E := \Big(\sum_{\text{edges}} \|X_\varepsilon - X_{\varepsilon[i]}\|_X^2 \Big)^{1/2} \leq (n2^{n-1})^{1/2}(1/n) = \frac{2^{n/2}}{\sqrt{2n}}.$$

If X had a metric type $p > 1$, then by the definition [B-M-W],

$$D \leq \alpha n^{1/p-1/2} E$$

for some constant α independent of n. But this would clearly fail for n large enough. Finally, since X has no metric type $p > 1$, it has no type $p > 1$ [B-M-W]. This contradiction completes the proof.

In particular, C_2^n cannot be semi-Lipschitz embedded into l_p^n ($1 < p < \infty$). However, we have

Proposition 8. *There is a sequence of mappings $F_n : C_2^n \to l_p$ such that*

(i) The family (F_n) is equicontinuous;
(ii) $\sup_n \|F_n^{-1}\|_{Lip} < \infty$.

Proof. Define F_n by

$$F_n(x) = n^{-1/p} x \quad \text{for } x \in C_2^n.$$

It is straightforward to verify that F_n^{-1} is 1-Lipschitz. Now pick any $x, y \in C_2^n$. Since the coordinates of x, y are either zero or one, we have

$$\|F(x) - F(y)\| = n^{-1/p} \|x - y\|_p = (d_n(x, y))^{1/p}.$$

This proves (i) and therefore completes the proof.

4. Embeddings into ℓ_∞^k.

We begin with a result stating that a Lipschitz map from a Lèvy family into l_∞^k concentrates. Given an k-dimensional normed space X, we denote by $d_\infty(X)$ its Banach-Mazur distance to l_∞^k.

Define the function

$$\varphi(k, n) = \sqrt{\frac{\ln(2 + k/n)}{n}}.$$

Theorem 9. *Let (T_n, d_n, μ_n) be a Lèvy family with constant c. Let (X_n) be a sequence of finite dimensional normed spaces with*

$$d_\infty(X_n) \varphi(\dim X_n, n) \to 0 \quad \text{as } n \to \infty. \tag{10}$$

Let $\varepsilon > 0$. If n is large enough then every 1-Lipschitz map $F : T_n \to X_n$ concentrates. That is,

$$\mathbf{P}\{\|F - \mathbb{E}F\|_{X_n} > \varepsilon\} \leq 8 \exp(-\frac{c}{16} \varepsilon^2 n).$$

The words "large enough" in the statement mean that there is a number n_0 which depends only on c, ε, and the rate of convergence in (10) such that the conclusion holds for every $n > n_0$.

In general, n_0 must depend on ε, as the following example shows. Let, for every n, $\varepsilon_n = (\log n)^{-1/2}$, $T_n = S^{n-1}$, and $X_n = l_{2+\varepsilon_n}^n$. Consider

$F_n : T_n \to X_n$, the formal identity in $\mathbf{R}^n$ restricted to S^{n-1}. Then (10) is satisfied:

$$d_\infty(X_n)\varphi(n,n) \approx n^{1/(2+\varepsilon_n)} n^{-1/2} \to 0 \quad (n \to \infty).$$

Assume that n_0 from the preceding discussion is independent of ε, and take $\varepsilon = n^{-1/4}$. Then F must concentrate for large n, in particular $\|F_n\|_{X_n} \le n^{-1/4}$ with non-zero probability provided n is large enough. This yields with non-zero probability

$$\|F_n\|_{l_2^n} \le n^{-1/4} d(l_2^n, X_n) = n^{-1/4} n^{1/2 - 1/(2+\varepsilon_n)} < n^{-1/8} \quad \text{for large } n.$$

This clearly contradicts the definition of F.

We postpone the proof of Theoremcorconc to the end of the section, and discuss some of its conclusions. First, Theorem 9 applies to Lipschitz embeddings.

Theorem 10. *Let (T_n, d_n, μ_n) be a regular Lèvy family. Then any map F from T_n into l_∞^k satisfies*

$$\|F\|_{Lip} \|F^{-1}|_{F(T_n)}\|_{Lip} \ge C(c, \varepsilon, \delta) \varphi(k,n)^{-1}, \tag{11}$$

where c, ε, δ are the constants from the definition of the regular Levy family.

The regularity assumption is essential here. Indeed, let T_n be any subset of l_∞^n. We consider it as a metric probability space with the induced metric and the probability concentrated in one (any) atom. Then (T_n, d_n, μ_n) is a Lèvy family and it is nicely embedded into l_∞^n.

There is an important instance of Theorem 10, namely when $T_n = S^{n-1}$ and F is a linear operator. This corresponds to the problem of approximating the standard Euclidean ball by a polyhedron with a given number of faces (or, alternatively, vertices). The precise estimate was obtained by Bourgain, Lindenstrauss and Milman [B-L-M], Carl and Pajor [C-P] and Gluskin [G 89]. Their result now follows from Theorem 10.

Corollary 11. *Let X be an n-dimensional subspace of l_∞^k. Then $d(X, l_2^n) \ge C\varphi(k,n)^{-1}$, where c is an absolute constant.*

Equivalently, let X be an n-dimensional normed space whose unit ball has at most k extreme points. Then $d(X, l_2^n) \ge C\varphi(k,n)^{-1}$.

The estimates are exact up to absolute constants, see [G 89], [G 86]. Moreover, *any* k-dimensional normed space Y has an n-dimensional subspace X, $n \ge \log k$, satisfying $d(X, l_2^n) \le C\varphi(k,n)^{-1}$ [M-S 98].

Our approach to (11) generalizes Corollary 11 to spaces different from l_2^n. Actually, the use of concentration of measure provides a result more general than Corollary 11. Recall a result of M. Gromov and V. Milman ([Gr-M], see also [Schm]), which states that the concentration of measure phenomenon holds on the unit spheres of uniformly convex spaces. Let X be a uniformly convex space. By a result of G. Pisier [Pi 75], there exists an equivalent norm on X such that the modulus of uniform convexity $\delta_X(t)$ grows as t^p for some $p > 1$. Gromov and Milman proved that there exists a (natural) probability measure on the unit sphere S_X of the space X such that any 1-Lipschitz function $F : S_X \to \mathbf{R}$ concentrates:

$$\mathbf{P}\{|F - \mathbb{E}F| > \varepsilon\} \leq 4\exp(-c\varepsilon^p n) \qquad \text{for } \varepsilon > 0.$$

Then a trivial modification of the technique used for Lèvy families iproves the following generalization of Corolary 11.

Corollary 12. *Let X be an n-dimensional subspace of l_∞^k, and let Y be a k-dimensional space whose modulus of uniform convexity satisfies*

$$\delta_Y(t) \leq Kt^p, \qquad t > 0.$$

Then $d(X,Y) \geq c(n/\log(2 + k/n))^{1/p}$, where c depends on K only.

The key to the results of this section is provided by the following lemma.

Lemma 13. *Let (T_n, d_n, μ_n) be a Lèvy family with constant c, and let $\varepsilon > 0$. Suppose $F : T_n \to l_\infty^k$ is a 1-Lipschitz function. Then there is a set $A \subset T_n$ such that*

$$\mu_n(A) \geq \frac{1}{2}\exp(-\varepsilon^2 n) \quad \text{and} \quad \operatorname{diam}(F(A)) \leq C(c,\varepsilon)\varphi(k,n).$$

Proof. One can assume that $k/n \geq e$ by embedding l_∞^k into some $l_\infty^{k'}$ with some larger k'. So we can substitute $\varphi(k,n)$ by $\sqrt{\ln(k/n)/n}$ in the statement of the Lemma. Let $t = t(c,\varepsilon) > 0$ be a number to be defined later. Write $F = (f_1, \dots, f_k)$, where all f_i are real valued 1-Lipschitz functions; we can also assume that they are mean zero.

Define a map $T : T_n \to \mathbb{Z}^k$ as follows: $T(\omega) = (b_1(\omega), \dots, b_k(\omega))$, where

$$b_i(\omega) \text{ is the nearest integer to } \frac{f_i(w)}{t\varphi(k,n)}.$$

thus

$$b_i(\omega) = s \quad \iff \quad s - 1/2 < \frac{f_i(\omega)}{t\varphi(k,n)} \leq s + 1/2.$$

Then for every i and $s \in \mathbb{Z}_+$

$$\begin{aligned} \mu_n\{|b_i(\omega)| \geq s\} &= \mu_n\Big\{s - 1/2 < \frac{f_i(\omega)}{t\varphi(k,n)} \text{ or } \frac{f_i(\omega)}{t\varphi(k,n)} \leq -s + 1/2\Big\} \\ &\leq \mu_n\{|f_i(\omega)| \geq (s - 1/2)t\varphi(k,n)\} \\ &\leq 8\exp\big(-c(s-1/2)^2 t^2 \varphi(k,n)^2 n\big) =: P, \end{aligned}$$

by (4). Since expectation is linear,

$$\mathbb{E}|\{i : |b_i(\omega)| \geq s\}| \leq kP.$$

Set

$$\alpha_s = P \cdot 2^{s+1}.$$

By Chebyshev's inequality

$$\mu_n\Big\{|\{i : |b_i(x)| \geq s\}| \geq k\alpha_s\Big\} \leq 1/2^{s+1}.$$

Now we define a set $B \subset \mathbb{Z}^n$ by

$$(b_1, \dots, b_n) \in B \quad \Longleftrightarrow \quad |\{i : |b_i| \geq s\}| \leq k\alpha_s \text{ for all } s \in \mathbb{Z}_+.$$

Then

$$\mu_n\{Tx \in B\} \geq 1 - \sum_{s=1}^{\infty} 1/2^{s+1} = 1/2. \tag{12}$$

CLAIM. $|B| \leq \exp(\varepsilon^2 n)$.

Once this is proved, we can apply the pigeonhole principle to (12). There exists a set $A \subset T_n$ with $\mu_n(A) \geq \frac{1}{2}\exp(-\varepsilon^2 n)$ such that $T(A)$ is a singleton. This will clearly complete the proof of lemma.

By the definition,

$$\alpha_s = 8 \cdot 2^{s+1}(k/n)^{-c(s-1/2)^2 t^2}.$$

Now we proceed by a counting argument from [Sp]. By taking t large enough we can assume that $1/2 > \alpha_1 > \alpha_2 > \dots$ Then

$$|B| \leq \prod_{s=1}^{\infty} \Big[\Big(\sum_{i=0}^{k\alpha_s} \binom{k}{i}\Big) 2^{k\alpha_s}\Big].$$

Indeed, $\{i : |b_i| = s\}$ can be chosen in at most $\sum_{i=0}^{k\alpha_s} \binom{k}{i}$ ways, and, having been selected, can be split into $\{i : b_i = s\}$ and $\{i : b_i = -s\}$ in at most $2^{k\alpha_s}$ ways. We bound

$$\sum_{i=0}^{k\alpha_s} \binom{k}{i} \leq 2^{kH(\alpha_s)},$$

where $H(\alpha) = -\alpha \log_2 \alpha - (1-\alpha)\log_2(1-\alpha)$, see [Ch]. Therefore

$$|B| \leq 2^{bk}, \quad \text{where } b = \sum_{s=1}^{\infty} (H(\alpha_s) + \alpha_s). \tag{13}$$

Note that

$$\alpha_{s+1} \leq \gamma\alpha_s \quad \text{and} \quad \alpha_s \leq \alpha_1 \leq \gamma \text{ for all } s \geq 1,$$

where

$$\gamma = 32\exp(-ct^2/4) \leq 1/100$$

(by an appropriate choice of t). Then

$$H(\alpha_{s+1}) + a_{s+1} \leq 3\gamma(H(\alpha_s) + a_s) \quad \text{for all } s \geq 1.$$

This yields that the series for b in (13) is dominated by the first term:

$$\begin{aligned} b &\leq \frac{H(\alpha_1)+\alpha_1}{1-3\gamma} \\ &\leq -3\alpha_1 \log_2(\alpha_1) \qquad \text{since } \alpha_1 \leq \gamma \leq 1/100, \\ &= 24ct^2(k/n)^{-ct^2/4}\log_2(32k/n). \end{aligned} \tag{14}$$

Then $|B| \leq \exp(\beta n)$, where $\beta = (\ln 2)(k/n)b$. If $t = t(c,\varepsilon)$ is chosen large enough, then $\beta \leq \varepsilon$, since $k/n \geq e$. This proves the Claim and completes the proof of the lemma.

The following simple lemma was already used by N. Alon and V. Milman [Al-M]. Recall that $\alpha(T,\varepsilon)$ is the concentration function of T defined by (2).

Lemma 14. *Let (T,d,μ) be a metric probability space. If $A \subset T$ with $\mu(A) \geq \alpha(T,\varepsilon/2)$, then $\mu(A_\varepsilon) \geq 1-\alpha(T,\varepsilon/2)$.*

In particular, for any subset $A \subset T$

$$\mu(A_\varepsilon) \geq 1 - \frac{\alpha(T,\varepsilon/2)}{\mu(A)}.$$

Proof. Let $\mu(A) \geq \alpha(T,\varepsilon/2)$. We claim that $\mu(A_{\varepsilon/2}) \geq 1/2$. Assume the converse. That is, assume $\mu(A_{\varepsilon/2})^c > 1/2$. Then $\mu((A_{\varepsilon/2})^c)_{\varepsilon/2} \geq 1-\alpha(T,\varepsilon/2)$. Clearly, $((A_{\varepsilon/2})^c)_{\varepsilon/2} \cap A = \emptyset$, thus $\mu(A) \leq \alpha(T,\varepsilon/2)$. This contradicts the assumption and proves the claim. Now

$$\mu(A_\varepsilon) = \mu(A_{\varepsilon/2})_{\varepsilon/2} \geq 1-\alpha(T,\varepsilon/2).$$

The second statement of the Lemma follows from the first one. The proof is complete.

Theorem 15. *Let (T_n,d_n,μ_n) be a Lèvy family with constant c, and let $\varepsilon > 0$. Let X be a k-dimensional Banach space, and $F : T_n \to X$ be a 1-Lipschitz map. Then*

$$\mu_n\Big\{\|F-\mathbb{E}F\| > \varepsilon + C(c,\varepsilon)d_\infty(X)\varphi(k,n)\Big\} \leq 8c_1\exp(-\frac{c}{16}\varepsilon^2 n). \tag{15}$$

Proof. Let $c_1 = C(\varepsilon,c)$. One can assume that $8\exp(-\frac{c}{16}\varepsilon^2 n) \leq 1$. We get a set A from Lemma 13 so that

$$\mu_n(A) \geq \frac{1}{2}\exp(-c(\varepsilon/4)^2 n) \;\text{ and }\; \operatorname{diam}(F(A)) \leq \frac{c_1}{2}d_\infty(X)\varphi(k,n).$$

Then $\mu_n(A) \geq \alpha(T_n, \varepsilon/4)$. Lemma 14 gives for $\delta \geq \varepsilon$

$$\mu_n(A_{\delta/2}) \geq 1 - \alpha(T_n, \delta/4) \geq 1 - 8\exp(-\frac{c}{16}\delta^2 n). \tag{16}$$

As F is 1-Lipschitz, it stabilizes not only on A but also on A_δ, so that we have for all $\delta \geq \varepsilon$ and for all $\omega \in A_{\varepsilon/2} \subset A_{\delta/2}$

$$\begin{aligned} \mu_n\Big\{\omega' : \|F(\omega) - F(\omega')\|_X > \delta + \frac{c_1}{2} d_\infty(X)\varphi(k,n)\Big\} &\leq \mu_n(A_{\delta/2})^c \\ &\leq 8\exp(-\frac{c}{16}\delta^2 n). \end{aligned}$$

Then for every $x \in A_{\varepsilon/2}$ we bound

$$\begin{aligned} \|F(\omega) - \mathbb{E}F\|_X &\leq \int_{T_n} \|F(\omega) - F(\omega')\|_X \, d\mu_n(\omega') \\ &\leq \varepsilon + \frac{c_1}{2} d_\infty(X)\varphi(k,n) + \int_\varepsilon^\infty 8\exp(-\frac{c}{16}\delta^2 n)\, d\delta \\ &\leq \varepsilon + \frac{c_1}{2} d_\infty(X)\varphi(k,n) + 8\sqrt{\frac{\pi}{4Cn}} \\ &\leq \varepsilon + c_1 d_\infty(X)\varphi(k,n) \quad \text{by adjusting } c_1. \end{aligned}$$

then the measure in (15) does not exceed $\mu_n(A_{\varepsilon/2})^c$ which, in turn, is majorized by (16). This concludes the proof.

Now Theorem 9 follows immediately from Theorem 15.

Proof of Theorem 10. We can assume that $\|F\|_{Lip} = 1$.

Choose δ and ε from the definition of regularity. Let $\delta_0 = \delta/2$. Clearly, we may assume that $n > n_0$, where $n_0 = n_0(c, \delta, \varepsilon)$ is large enough. Then we take a set A given by Lemma 13 so that

$$\mu_n(A) \geq \alpha(T_n, \delta_0/2)$$

and

$$\mu_n(A) \geq \alpha(T_n, \delta_0/2) \quad \text{and} \quad \mathrm{diam}(F(A)) \leq C(c,\delta)\varphi(k,n). \tag{17}$$

Then Lemma 14 yields

$$\begin{aligned} \mu_n(A_{\delta_0}) &\geq 1 - \alpha(T_n, \delta_0/2) \\ &\geq 1 - 4\exp(-(c/4)\delta_0^2 n) \geq 1 - \varepsilon \end{aligned}$$

provided n_0 was chosen sufficiently large, and $n > n_0$. Then we get from the regularity of (T_n, d_n, μ_n) that $\mathrm{diam}(A_{\delta_0}) \geq \delta$. Thus $\mathrm{diam}(A) \geq \delta - \delta_0 = \delta/2$. Together with (17) this gives $\|F^{-1}|_{F(T_n)}\|_{Lip} \geq (\delta/2)(C(c,\delta)\varphi(k,n))^{-1}$, completing the proof.

References

[Al-M] N. Alon, V. Milman, *Embedding of l_k^∞ in finite-dimensional Banach spaces*, Israel J. Math. 45 (1983),265–280

[Bo] J. Bourgain, *On Lipschitz embeddings of finite metric spaces in Hilbert space*, Israel J. Math. 52 (1992), 103–111

[B-L-M] J. Bourgain, J. Lindenstrauss, V. Milman, *Approximation of zonoids by zonotopes*, Acta Math. 162 (1989), 73–141

[B-M-W] J. Bourgain, V. Milman, H. Wolfson, *On type of metric spaces*, Trans. AMS 294 (1986), 295–317

[C-P] B. Carl, A. Pajor, *Gelfand numbers of operators with values in Hilbert spaces*, Invent. Math. 94 (1988), 79–103

[Ch] H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statist. 23 (1952), 493–509

[G 86] E. Gluskin, *An octahedron is poorly approximated by random subspaces*, Funktsional. Anal. i Prilozhen. 20 (1986), no. 1, 14–20; English transl. in Functional Anal. Appl. 20 (1986)

[G 89] E. Gluskin, *Extremal properties of orthogonal parallelepipeds and their applications to the geometry of Banach spaces*, Math. USSR Sbornik 64 (1989), 85–96

[G-M-R] A. A. Giannopoulos, V. D. Milman, M. Rudelson, *Convex bodies with minimal mean width*, to appear in GAFA

[Gr-M] M. Gromov, V. Milman, *Generalization of the spherical isoperimetric inequality to uniformly convex Banach spaces*, Compositio Math. 62 (1987), 263–282

[J-L-S] W. Johnson, J. Lindenstrauss, G. Schechtman, *On Lipschitz embedding of finite metric spaces in low dimensional normed spaces*, in Geometrical Aspects of Functional Analysis (J. Lindenstrauss and V. Milman, eds), Lecture Notes in Math. 1267, Springer-Verlag, 1987, 1774-184

[Le-Ta] M. Ledoux, M. Talagrand, *Probability in Banach spaces*, A Series of Modern Surveys in Math., 23, Springer Verlag, 1991.

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

[Ma 95] J. Matoušek, *On the distortion required for embedding finite metric spaces into normed spaces*, Israel J. Math. 93 (1996), 333–344

[Ma 97] J. Matoušek, *On embedding expanders into ℓ_p spaces*, Israel J. Math. 102 (1997), 189–197

[Ma 99] J. Matoušek, *On embedding trees into uniformly convex Banach spaces*, Israel J. Math. 114 (1999), 221–237

[M-S] V. Milman, G. Schechtman, *Asymptotic theory of finite dimensional normed spaces*, Lecture Notes in Math., vol. 1200, Springer Verlag, 1986.

[M-S 98] V. Milman and G. Schechtman, *An "isomorphic" version of Dvoretzky's theorem, II*, to appear in Convex Geometry, MSRI Publications 34 (1998)

[Pi 75] G. Pisier, *Martingales with values in uniformly convex spaces*, Israel J. Math. 20 (1975), 326–350

[Pi 89] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Tracts in Math. 94. Cambridge University Press, Cambridge, 1989

[Schm] M. Schmuckenschlaeger, *A concentration of measure phenomenon on uniformly convex bodies*, Operator Theory, Advances and Appl. 70 (1995), 275–287

[Sp] J. Spencer, *Six standard deviations suffice*, Trans. AMS 289 (1985), 679–706

[Ta] M. Talagrand, *A new look at independence*, Annals of Probability 24 (1996), 1–34

E-mail address: rudelson@math.missouri.edu, vershyn@wisdom.weizmann.ac.il