Random $\varepsilon$ nets and embeddings in $\ell^N_\infty$

Y. Gordon† A. E. Litvak A. Pajor
N. Tomczak-Jaegermann ‡

Abstract

We show that, given an $n$-dimensional normed space $X$ a sequence of $N = (8/\varepsilon)^{2n}$ independent random vectors $(X_i)_{i=1}^N$, uniformly distributed in the unit ball of $X^*$, with high probability forms an $\varepsilon$ net for this unit ball. Thus the random linear map $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$ defined by $\Gamma x = (\langle x, X_i \rangle)_{i=1}^N$ embeds $X$ in $\ell^N_\infty$ with at most $1 + \varepsilon$ norm distortion. In the case $X = \ell^n_2$ we obtain a random $1 + \varepsilon$ embedding into $\ell^N_\infty$ with asymptotically best possible relation between $N$, $n$, and $\varepsilon$.

1 Introduction

Let $X = (\mathbb{R}^n, \| \cdot \|)$ be an arbitrary $n$-dimensional normed space with unit ball $K$. It is well known that, for any $0 < \varepsilon < 1$, $X$ can be $1 + \varepsilon$ embedded into $\ell^N_\infty$, for some $N = N(\varepsilon, n)$, depending on $\varepsilon$ and $n$, but independent of $X$. In this note we investigate $1 + \varepsilon$ isomorphic embeddings which are random with respect to some natural measure, depending on $K$. We first show that for $N = (8/\varepsilon)^{2n}$, a sequence of $N$ independent random vectors $(X_i)_{i=1}^N$, uniformly distributed in the unit ball $K^0$ of the dual space $X^*$, forms an $\varepsilon$ net for $K^0$, with high probability. Thus, with high probability, the random linear map

---

*Keywords: Dvoretzky theorem, random embeddings into $\ell_\infty$, random $\varepsilon$-nets. 2000 Mathematical Subject Classification: 46B07, 46B09, 52A21, 60D05.
†Partially supported by France-Israel exchange fund 2005389, and by the Fund for the Promotion of Research at the Technion.
‡This author holds the Canada Research Chair in Geometric Analysis.
\( \Gamma : \mathbb{R}^n \to \mathbb{R}^N \) defined by \( \Gamma x = (\langle x, X_i \rangle)_{i=1}^N \) embeds \( X \) in \( \ell_\infty^N \) with at most \( 1 + \varepsilon \) norm distortion.

The important case is \( X = \ell_2^n \). In this case it is more natural to consider random vectors \( X_i \) uniformly distributed on the sphere \( S^{n-1} \). Such vectors also form an \( \varepsilon \)-net on the sphere hence they determine a random \( 1 + \varepsilon \) embedding \( \Gamma \) of \( \ell_2^n \) into \( \ell_\infty^N \). We also show that \( \sqrt{n/N} \Gamma \) is a \( 1 + \varepsilon \) isometry from \( \ell_2^n \) into \( \ell_\infty^N \), with high probability.

The case \( X = \ell_2^n \) is connected with Dvoretzky’s theorem ([D]). Milman found a new proof ([M]), using the Levy isoperimetric inequality on the sphere, that there exists a function \( c(\varepsilon) > 0 \) such that for all \( n \leq c(\varepsilon) \log N \), \( \ell_2^n \) can be \( 1 + \varepsilon \) embedded into any normed space \( Y \) of dimension \( N \). His proof gives \( c(\varepsilon) \sim \varepsilon^2 / \log(2/\varepsilon) \). Later a new approach was found in ([Go1]) by using random Gaussian embeddings. It yields that \( c(\varepsilon) \sim \varepsilon^2 \) is sufficient. Milman raised the question what is the best behavior of \( c(\varepsilon) \), as \( \varepsilon \to 0 \), in the above estimates. Recently Schechtman showed in [S1] that one may take \( c(\varepsilon) \sim \varepsilon / (\log(2/\varepsilon))^2 \), however his approach is not random.

Since in this paper we deal with embeddings into \( \ell_\infty^N \), we shall restrict our attention to this case only. When \( Y = \ell_\infty^N \), it is well known that there exists an embedding with \( c(\varepsilon) \sim 1 / \log(2/\varepsilon) \). It is also known that this behavior of \( c(\varepsilon) \) as \( \varepsilon \to 0 \) cannot be improved. The standard embedding relies on the existence of \( \varepsilon \)-nets of appropriate cardinalities. It is therefore natural to ask whether this embedding can be randomized.

In this paper we provide a positive answer to this question. Namely, we show (in Theorems 4.1, 4.3) that for the random embedding \( \Gamma \) determined by independent uniformly distributed vectors on \( S^{n-1} \), with large probability one may achieve \( c(\varepsilon) \sim \frac{1}{\log(2/\varepsilon)} \), which is the best possible as mentioned above. We would like to note that such result is not valid in the setting of the Haar measure on Grassman manifold (equivalently, for embeddings defined by Gaussian matrices). Indeed, Schechtman recently showed ([S2]) that if “most” \( n = c'(\varepsilon) \log N \) dimensional subspaces of \( \ell_\infty^N \) are \( 1 + \varepsilon \) Euclidean then \( c'(\varepsilon) \sim \varepsilon \).

2 Notation and preliminary results

We denote by \( \langle \cdot, \cdot \rangle \) the scalar product of the canonical Euclidean structure on \( \mathbb{R}^n \) and by \( | \cdot | \) the canonical Euclidean norm. The Euclidean ball is denoted by \( B_2^n \) and the Euclidean sphere is denoted by \( S^{n-1} \).
By a convex body in \( \mathbb{R}^n \) we always mean a compact convex set with non-empty interior. A centrally symmetric body with respect to origin will be called symmetric.

Given a convex body \( K \) in \( \mathbb{R}^n \) we denote by \( |K| \) its volume and by \( \| \cdot \|_K \) we denote the Minkowski functional of \( K \), i.e.

\[
\|x\|_K = \inf \{ \lambda > 0 \mid x \in \lambda K \}.
\]

If \( K \) is symmetric then \( \| \cdot \|_K \) is a norm with the unit ball \( K \).

Given a finite set \( A \) we denote its cardinality by \( |A| \).

Recall that if \( K \) is a symmetric convex body in \( \mathbb{R}^n \) then for every \( 0 < \varepsilon \leq 1 \) there exists an \( \varepsilon \)-net \( \Lambda \) in \( K \) with respect to the norm \( \| \cdot \|_K \) of cardinality

\[
|\Lambda| \leq (1 + 2/\varepsilon)^n \leq (3/\varepsilon)^n.
\]

The polar of a convex body \( K \subset \mathbb{R}^n \) is defined by

\[
K^0 = \{ x \mid \langle x, y \rangle \leq 1 \text{ for every } y \in K \}.
\]

Let \( K \) be a convex body. We say that a vector \( X \) is uniformly distributed on \( K \) if

\[
P(\{ X \in A \}) = \frac{|K \cap A|}{|K|}
\]

for every measurable \( A \subset \mathbb{R}^n \).

Given a square matrix \( T \) by \( \| T \|_{HS} \) we denote its Hilbert-Schmidt norm.

Below we will need the following geometric lemma. Although we will use only a particular case of the lemma, we prefer to state it in full generality for future references.

**Lemma 2.1** Let \( d > 0 \) and \( K, L \) be convex bodies in \( \mathbb{R}^n \) such that \( K \subset -dL \).

Then for every \( x \in K \) and for \( 0 < \varepsilon \leq 1 \) one has

\[
|K \cap (x + \varepsilon L)| \geq \left| \frac{\varepsilon}{d+1} K \cap L \right|.
\]

In particular, if \( K = L = -K \) then

\[
|K \cap (x + \varepsilon K)| \geq \left| \frac{\varepsilon}{2} K \right|.
\]
Proof: Denote 
\[ \alpha = 1 - \frac{\epsilon}{d+1}, \quad \beta = \frac{\epsilon}{d+1}. \]
To prove the desired result it is enough to show that 
\[ K \cap (x + \epsilon L) \supset \alpha x + \beta K \cap L. \]
Let \( z = \alpha x + \beta y \), where \( y \in K \cap L \). Clearly, \( z \in K \) and \( z = x + \beta(y - x) \).
Since 
\[ y - x \in L - K \subset L + dL = (1 + d)L, \]
we obtain the result. \( \square \)

Remark 1. The example of the cube (when \( x \) is a vertex) shows that the estimate in the “in particular” part of Lemma 2.1 is sharp.

Remark 2. It is known that for every convex body \( K \) in \( \mathbb{R}^n \) there exists a shift such that \( K - a \subset -n(K - a) \). Thus, Lemma 2.1 implies that for every convex body \( K \) in \( \mathbb{R}^n \) there exists a vector \( a \in \mathbb{R}^n \) such that for every \( x \in K \) and for \( \epsilon > 0 \) one has 
\[ |(K - a) \cap (x + \epsilon (K - a))| \geq \left| \frac{\epsilon}{n+1} K \right|. \]
The example of the regular simplex (when \( x \) is a vertex) shows that the latter estimate is sharp.

Remark 3. It was proved in [GLMP] that if a body \( L \) is in the position of maximal volume in \( K \) (that is \( L \subset K \) and for every linear map \( T \) and every point \( x \in \mathbb{R}^n \) satisfying \( TL + x \subset K \) one has \(|TL| \leq |L|\)), then there exist \( a \in \mathbb{R}^n \) such that 
\[ L - a \subset K - a \subset -n(L - a). \]
Thus Lemma 2.1 implies that if a body \( L \) is in the position of maximal volume in \( K \) then there exists a vector \( a \in \mathbb{R}^n \) such that for every \( x \in K \) and for \( \epsilon > 0 \) one has 
\[ |(K - a) \cap (x + \epsilon (L - a))| \geq \left| \frac{\epsilon}{n+1} L \right|. \]
3 Random embeddings of normed spaces in $\ell^N_\infty$

First we show that $N$ vectors uniformly distributed on a symmetric convex body $K$ form an $\varepsilon$-net in $K$.

**Theorem 3.1** Let $n \geq 1$, $0 < \varepsilon \leq 1$, and $N = (4/\varepsilon)^{2n}$. Let $X_1, \ldots, X_N$ be independent random variables uniformly distributed on a symmetric convex body $K$ in $\mathbb{R}^n$. Then with a probability larger than $1 - \exp (- (8/\varepsilon)^n / 2)$ the set $\mathcal{N} = \{X_1, \ldots, X_N\}$ forms an $\varepsilon$-net in $K$.

**Proof:** Fix an $\varepsilon/2$-net $\Lambda \subset K$ with $|\Lambda| \leq (6/\varepsilon)^n$, and consider random vectors $X_1, \ldots, X_N$ uniformly distributed on $K$, where $N$ is as in the statement. We want to show that the probability

$$ P \{ \forall x \in K \exists i \leq N \text{ such that } \|x - X_i\|_K < \varepsilon \} \quad (1) $$

is large. Clearly this probability is larger than

$$ P \{ \forall x \in \Lambda \exists i \leq N \text{ such that } \|x - X_i\|_K < \varepsilon / 2 \} \quad (2) $$

By $A$ denote the event considered in (2), and estimate the probability of its complement $A^c$. We have

$$ P(A^c) = P \{ \exists x \in \Lambda \forall i \leq N \text{ one has } \|x - X_i\|_K \geq \varepsilon / 2 \} $$

$$ \leq |\Lambda| \left( P \{ \|x_0 - X_1\|_K \geq \varepsilon / 2 \} \right)^N $$

$$ \leq |\Lambda| \left( 1 - P \{ \|x_0 - X_1\|_K < \varepsilon / 2 \} \right)^N, $$

where $x_0 \in \Lambda$ satisfies

$$ P \{ \|x_0 - X_i\|_K \geq \varepsilon / 2 \} = \max_{x \in \Lambda} P \{ \|x - X_i\|_K \geq \varepsilon / 2 \}.$$

Note that

$$ P \{ \|x_0 - X_1\|_K < \varepsilon / 2 \} = P \left\{ X_1 \in x_0 + \frac{\varepsilon}{2} K \right\} = \frac{|K \cap (x_0 + \frac{\varepsilon}{2} K)|}{|K|}. $$

Applying Lemma 2.1 we obtain

$$ P \{ \|x_0 - X_i\|_K < \varepsilon / 2 \} \geq \left( \frac{\varepsilon^2}{2} \right)^n. $$
This implies
\[
\mathbb{P}(A^c) \leq (6/\varepsilon)^n \left(1 - (\varepsilon/2)^n\right)^N \leq (6/\varepsilon)^n \exp\left(-\frac{8}{\varepsilon} n^2\right)
\]
which implies the result.

To prove the next theorem we need the following standard lemma. We provide its proof for the sake of completeness.

**Lemma 3.2** Let \( X \) be a Banach space and \( K \) be its unit ball. Let \( \mathcal{N} \) be an \( \varepsilon \)-net in the unit ball \( K^0 \) (or in the unit sphere \( \partial K^0 \)) of the dual space. Then for every \( x \in X \) we have
\[
\sup_{y \in \mathcal{N}} \langle x, y \rangle \leq \|x\|_K \leq (1 - \varepsilon)^{-1} \sup_{y \in \mathcal{N}} \langle x, y \rangle.
\]

**Proof:** The left hand side estimate is obvious. Now let \( \|x\|_X = 1 \) and consider \( z \in \partial K^0 \) such that \( \langle x, z \rangle = 1 \). Then for an appropriate \( y \in \mathcal{N} \) we have
\[
1 = \langle x, y \rangle + \langle x, z - y \rangle \leq \sup_{y \in \mathcal{N}} \langle x, y \rangle + \varepsilon,
\]
which implies the required estimate.

Combining Theorem 3.1 with Lemma 3.2 we obtain that a random matrix whose rows are independent random vectors uniformly distributed on the polar of a symmetric convex body provides a random embedding of the body into \( \ell^\infty_N \).

**Theorem 3.3** Let \( 0 < \varepsilon < 1 \) and \( n \leq \frac{\log N}{2 \ln(4/\varepsilon)} \). Let \( K \) be a symmetric convex body in \( \mathbb{R}^n \). Let \( X_1, \ldots, X_N \) be independent random vectors uniformly distributed on \( K^0 \). Consider the matrix \( \Gamma : \mathbb{R}^n \to \mathbb{R}^N \) whose rows are \( X_1, \ldots, X_N \) (i.e. \( \Gamma x = \langle (x, X_i)\rangle_{i=1}^N \)). Then with probability larger than
\[
1 - \exp\left(-\frac{8}{\varepsilon} n^2\right)
\]
we have
\[
(1 - \varepsilon) \|x\|_K \leq \|\Gamma x\|_\infty \leq \|x\|_K,
\]
for all \( x \in \mathbb{R}^n \).
4 The Euclidean case

In this section we discuss the embedding of $\ell_2^n$ into $\ell_\infty^N$. Here it is more natural to work with random vectors uniformly distributed on the Euclidean sphere $S^{n-1}$. Accordingly, in the rest of the paper $X_1, \ldots, X_N$ stands for independent random vectors uniformly distributed on the Euclidean sphere $S^{n-1}$ and $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$ is the matrix whose rows are $X_1, \ldots, X_N$ (that is, $\Gamma x = (\langle x, X_i \rangle)_{i=1}^N$).

One can easily check that Theorem 3.1 holds for $S^{n-1}$ and such vectors $X_1, \ldots, X_N$. Indeed, this follows from the same argument as before with minor modifications. We need only to observe that given $y \in S^{n-1}$ the normalized Lebesgue measure of a cap

$$\{ x \in S^{n-1} \mid |x - y| \leq \varepsilon \}$$

is larger than or equal to $(\varepsilon/2)^n$ (cf. e.g., [P1], chapter 6), as well as the fact that in $S^{n-1}$ there exists an $\varepsilon$-net of the cardinality $(3/\varepsilon)^n$. Therefore Theorem 3.3 holds with $K = B^n_2$ and with the matrix $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$ defined above. We formally state both facts for future reference.

**Theorem 4.1** Let $0 < \varepsilon < 1$ and $n \leq \frac{\log N}{2 \ln(4/\varepsilon)}$. Let $\mathcal{N} = \{X_1, \ldots, X_N\}$ where $X_i$ ($i = 1, \ldots, N$) are independent random vectors uniformly distributed on $S^{n-1}$. Then, with probability larger than $1 - \exp(-8/\varepsilon)^n/2$, $\mathcal{N}$ forms an $\varepsilon$-net on $S^{n-1}$. Furthermore, with the same probability, the matrix $\Gamma$ as above satisfies

$$(1 - \varepsilon) |x| \leq \|\Gamma x\|_\infty \leq |x|,$$

for all $x \in \mathbb{R}^n$.

Denote by $Q$ the unit ball of $\ell_\infty^N$ (i.e. the $N$-dimensional cube). Theorem 4.1 shows that $Q$ has an $n$-dimensional section (which can be realized as $E := \Gamma \mathbb{R}^n$) which is almost Euclidean, i.e.

$$\Gamma B^n_2 \subset Q \cap E \subset (1 - \varepsilon)^{-1} \Gamma B^n_2.$$

Below we show that in fact the ellipsoid $\Gamma B^n_2$ is, up to $\frac{1+\varepsilon}{1-\varepsilon}$, equivalent to the standard Euclidean ball of radius $\sqrt{N/n}$. In other words, a random subspace $E = \Gamma \mathbb{R}^n$ of $\ell_\infty^N$ is nearly Euclidean with respect to the canonical Euclidean structure on $\mathbb{R}^N$. Namely, Theorem 4.3 below shows that

$$\left(1 - \varepsilon\right) \sqrt{N/n} \Gamma B^n_2 \subset Q \cap E \subset \frac{1+\varepsilon}{1-\varepsilon} \sqrt{N/n} \Gamma B^n_2.$$
We need the following lemma, which shows that \( \sqrt{n/N} \Gamma \) almost preserves the Euclidean norm of a vector.

**Lemma 4.2** Let \( 0 < \varepsilon < 1 \) and let \( N \geq n^3/\varepsilon^4 \). Let \( X_1, \ldots, X_N \) be independent random points on the sphere \( S^{n-1} \). Then with probability larger than \( 1 - n^2/(\varepsilon^4N) \) we have

\[
(1 - \varepsilon) |x| \leq |\Gamma x| \sqrt{n/N} \leq (1 + \varepsilon) |x|,
\]

for every \( x \in \mathbb{R}^n \).

**Remark.** One can get better estimates using a theorem of Bourgain [B]. For instance, the above inequalities are satisfied with probability larger than \( 1 - \delta \) as far as \( N \geq c(\delta)n(\log n)^3/\varepsilon^2 \) (instead of let \( N \geq n^3/\varepsilon^4 \)) for some function \( c(\delta) > 0 \). However, we prefer to present here a simpler proof, which provides estimates good enough for our purposes.

**Proof:** Set \( A := \|\Gamma^*\Gamma - (N/n)I\|_{HS} \). Using the fact that \( \|T\|_{HS}^2 = \text{tr}(T^*T) \) for every operator \( T : \mathbb{R}^n \to \mathbb{R}^N \), we get

\[
A^2 = \sum_{i,j} |\langle X_i, X_j \rangle|^2 + (N^2/n^2)n - (2N/n)\|\Gamma\|_{HS}^2
\]

\[
= \sum_{i,j} |\langle X_i, X_j \rangle|^2 + (N^2/n^2)n - 2N^2/n
\]

\[
= \sum_i |X_i|^4 + \sum_{i \neq j} |\langle X_i, X_j \rangle|^2 - N^2/n.
\]

Therefore,

\[
\mathbb{E}A^2 = N + N(N - 1)\mathbb{E}|\langle X_1, X_2 \rangle|^2
\]

Since \( \mathbb{E}|\langle X_1, X_2 \rangle|^2 = 1/n \), we finally obtain \( \mathbb{E}A^2 = N(1 - 1/n) \).

By Chebyshev’s inequality we get, for any \( \varepsilon_1 > 0 \),

\[
\mathbb{P}\{A > \varepsilon_1\} \leq \mathbb{E}A^2/\varepsilon_1^2 \leq N/\varepsilon_1^2.
\]

Thus

\[
\mathbb{P}\left\{\left\| \frac{n}{N} \Gamma^*\Gamma - I \right\| < \varepsilon_1 \right\} \geq \mathbb{P}\left\{\left\| \frac{n}{N} \Gamma^*\Gamma - I \right\|_{HS} < \varepsilon_1 \right\}
\]

\[
\geq 1 - \mathbb{P}\{A > \frac{N}{n}\varepsilon_1\} \geq 1 - \frac{n^2}{\varepsilon_1^2 N^2} = 1 - \frac{n^2}{\varepsilon_1^2 N}.
\]
The last estimate implies that, for any $\varepsilon_1 > 0$, with probability larger than or equal to $1 - n^2/(\varepsilon_1^2 N)$, we have the estimates for singular numbers of the matrix $\Gamma$,
\[
\left| \sqrt{n/N}s_j(\Gamma) - 1 \right| < \sqrt{\varepsilon_1},
\]
for $j = 1, \ldots, n$. In particular,
\[
1 - \sqrt{\varepsilon_1} < \sqrt{n/N}s_n(\Gamma) \leq \sqrt{n/N}s_1(\Gamma) < 1 + \sqrt{\varepsilon_1}.
\]
Setting $\varepsilon_1 = \varepsilon_2$ immediately implies the desired conclusion.

Combining Theorem 3.3 with Lemma 4.2 we obtain the following theorem.

**Theorem 4.3** Let $0 < \varepsilon < 1$ and $2 \leq n \leq \log N / \ln(4/\varepsilon)$. Let $X_1, \ldots, X_N$ be independent random vectors uniformly distributed on $S^{n-1}$. Consider the matrix $\Gamma : \mathbb{R}^n \to \mathbb{R}^N$ whose rows are $X_1, \ldots, X_N$. Then with probability larger than $1 - n^2 \varepsilon^{2n-4}/16n - \exp(-(8/\varepsilon)n/2) \geq 1 - e^{-n}$ we have
\[
1 - \varepsilon |\Gamma x| \leq \sqrt{n/N} \|\Gamma x\|_{\infty} \leq \frac{1}{1 - \varepsilon} |\Gamma x|,
\]
for all $x \in \mathbb{R}^n$.

Finally, we would like to emphasize the differences between the randomness given by the matrix $\Gamma$ and a standard Gaussian matrix $G$ (i.e., with independent $N(0, 1)$ entries). Fix $N$ and $0 < \varepsilon < 1$. As already mentioned in the introduction, $\Gamma$ gives a random embedding with $n_1 \sim \log N / \log(2/\varepsilon)$ (which is best possible in general), while $G$ provides a random embedding with $n_2 \sim \varepsilon \log N$, which is best possible if one requires high probability ([S2]).

Another observation is that Euclidean sections of the cube determined by $\Gamma$ and $G$, and taken in the appropriate dimensions $n_1$ and $n_2$ (or smaller), will have different radii. Indeed, the conclusion of Theorem 4.3 implies that, with high probability defined by $\Gamma$, for every non-zero $y \in \Gamma \mathbb{R}^{n_1}$,
\[
\frac{\|y\|_{\infty}}{|y|} \sim \sqrt{\frac{n_1}{N}} \sim \sqrt{\frac{\log N}{N \log(2/\varepsilon)}}.
\]
On the other hand, with high probability defined by $G$ for every non-zero $y = Gx \in G \mathbb{R}^{n_2}$ one has
\[
\frac{\|y\|_{\infty}}{|y|} \sim \frac{\mathbb{E} \|Gx\|_{\infty}}{\mathbb{E} |Gx|} = \frac{\mathbb{E} \|Ge_1\|_{\infty}}{\mathbb{E} |Ge_1|} \sim \sqrt{\frac{\log N}{N}}.
\]
These two expectations are not comparable uniformly in $\varepsilon$ (as $\varepsilon \to 0$).

Added in the proof: Theorem 3.1 should be compared with Proposition 5.3 of [GiM]

References


Y. Gordon,
Department of Mathematics, Technion,
Haifa 32000, Israel,
gordon@technion.ac.il

A. E. Litvak and N. Tomczak-Jaegermann,
Department of Mathematical and Statistical Sciences,
University of Alberta,
Edmonton, Alberta, Canada T6G 2G1,
alexandr@math.ualberta.ca
nicole.tomczak@ualberta.ca

A. Pajor,
Équipe d’Analyse et Mathématiques Appliquées,
Université de Marne-la-Vallée,
5, boulevard Descartes, Champs sur Marne,
77454 Marne-la-Vallée, Cedex 2, France
Alain.Pajor@univ-mlv.fr