STUDIA MATHEMATICA 195 (3) (2009)

# Generalizing the Johnson–Lindenstrauss lemma to $k$-dimensional affine subspaces

by

ALON DMITRIYUK and YEHORAM GORDON (Haifa)

**Abstract.** Let $\varepsilon > 0$ and $1 \le k \le n$ and let $\{W_l\}_{l=1}^p$ be affine subspaces of $\mathbb{R}^n$, each of dimension at most $k$. Let $m = O(\varepsilon^{-2}(k + \log p))$ if $\varepsilon < 1$, and $m = O(k + \log p/\log(1+\varepsilon))$ if $\varepsilon \ge 1$. We prove that there is a linear map $H : \mathbb{R}^n \to \mathbb{R}^m$ such that for all $1 \le l \le p$ and $x, y \in W_l$ we have $\|x - y\|_2 \le \|H(x) - H(y)\|_2 \le (1 + \varepsilon)\|x - y\|_2$, i.e. the distance distortion is at most $1+\varepsilon$. The estimate on $m$ is tight in terms of $k$ and $p$ whenever $\varepsilon < 1$, and is tight on $\varepsilon, k, p$ whenever $\varepsilon \ge 1$. We extend these results to embeddings into general normed spaces $Y$.

**1. Introduction.** In 1984 Johnson and Lindenstrauss [17] proved the following important result, henceforth called the JL-lemma.

THEOREM 1.1. *Fix $0 < \varepsilon < 1$. Any $n$-point set in Euclidean space can be embedded into $O(\varepsilon^{-2} \ln n)$-dimensional Euclidean space with distance distortion smaller than $1 + \varepsilon$.*

Any embedding of the vertices of the standard $n$-simplex into $m$-dimensional Euclidean space with distortion $1 + \varepsilon$ satisfies $m \ge c(\varepsilon) \ln n$ where $c(\varepsilon)$ is a constant depending only on $\varepsilon$. Noga Alon [2] proved that $c(\varepsilon) \ge c\varepsilon^{-2}(\ln(1/\varepsilon))^{-1}$ where $c$ is an absolute constant.

The JL-lemma has been shown to be useful in applications in computer science and in engineering, and various proofs were given in [1, 3, 7, 10, 16].

We obtain some results which we then apply to *compressed sensing*, which is an important notion studied and applied by many authors (see e.g. [6, 8]). The relevant classical setup for compressed sensing is as follows: A set $T \subset \ell_2^n = (\mathbb{R}^n, \|\cdot\|_2)$ and an orthogonal basis $F = \{f_1, \dots, f_n\}$ are given. For $x \in \ell_2^n$ and $J \subset \{1, \dots, n\}$ we denote by $x_J$ the sparse vector in $\ell_2^n$ whose coordinates with respect to $F$ are zero outside of $J$ and are equal to the coordinates of $x$ with respect to $F$ on $J$. Given natural numbers

2010 *Mathematics Subject Classification*: 46B09, 46B07.
*Key words and phrases*: local theory, Gaussian processes, high dimensional geometry, convexity, normed linear spaces, Gaussian operators.

DOI: 10.4064/sm195-3-3  

$n > m > k$ we wish to know whether there is an embedding $G$ of $T$ into a smaller dimensional normed space $Y = (\mathbb{R}^m, \|\cdot\|_Y)$ such that for every $J \subset \{1,\dots,n\}$ with $|J| \le k$, the distance distortion of $G$ on $x_J$ is small. More precisely, given $\varepsilon > 0$ we seek an embedding $G$ of $T$ into $Y$ such that for every $J \subset \{1,\dots,n\}$ with $|J| \le k$,

$$(1-\varepsilon)\|x_J\|_2 \le \|G(x_J)\|_Y \le (1+\varepsilon)\|x_J\|_2.$$

This is easily obtained using Theorem 2.6 below by choosing the $\binom{n}{k}$ subspaces whose coordinates relative to $F$ are supported on $k$ elements. In the particular case of $Y = \ell_2^m$, we obtain a linear map $G$ of $\mathbb{R}^n$ into $\mathbb{R}^m$ where $m$ is of the order of magnitude $k\ln(n/k)$ or, equivalently, the order of magnitude of $k$ is $m/\ln(n/m)$. In fact Theorem 2.6 extends this to embeddings into a general Banach space $Y$, by using the Gaussian min-max theorem together with methods and tools applied in the asymptotic geometry of high dimensional convex bodies. In Theorem 2.3 we also improve and extend results obtained in [13, 14].

There are many applications of the compressed sensing theory and methods to reconstruct and recover digital signals and large amounts of sparse discrete information by embedding the data into a small dimensional space and maintaining very good $1+\varepsilon$ accuracy.

Our results generalize the JL-lemma in various directions including embeddings into general Banach spaces $Y$, and in particular give sharp results when $Y = \ell_2^m$. To do this we employ among other tools the Gaussian min-max theorem originally proved in [11] and extended in [12, 13]. The methods employed here have not been used in this context before, and we believe this new approach to the various extensions of the Johnson–Lindenstrauss lemma is important, as they usually provide sharp bounds and combine the probabilistic tools with other methods and tools employed in high dimensional asymptotic geometry. Though we use the Gaussian approach our methods and some theorems can be extended to any $\psi_2$ distributions, e.g., Bernouli random variables; we summarize how this may be done in the last section. It is important in applications, e.g. to compressed sensing, that a suitable random method provides the desired results, and in addition that with very high probability, a random selection yields the results. We prove this holds in all cases considered in this paper.

In the special case of embeddings between Hilbert spaces we obtain:

THEOREM 1.2. *Let $1 \le k \le n$ and let $\{W_l\}_{l=1}^p$ be subspaces of dimension at most $k$ in $\ell_2^n$.*

(i) *Let $0 < \varepsilon < 1$. Then there is a linear map $H : \ell_2^n \to \ell_2^m$ with $m = O(\varepsilon^{-2}(k + \ln p))$ such that for all $1 \le l \le p$ and $x, y \in W_l$,*

$$(1-\varepsilon)\|x-y\|_2 \le \|H(x)-H(y)\|_2 \le (1+\varepsilon)\|x-y\|_2.$$

(ii) *Let $D \ge 2$. Then there is a linear map $H : \ell_2^n \to \ell_2^m$ with $m = O(k + \ln p/\ln D)$ such that for all $1 \le l \le p$ and $x, y \in W_l$,*

$$\frac{1}{D}\|x-y\|_2 \le \|H(x)-H(y)\|_2 \le \|x-y\|_2.$$

(iii) *If $p \le \binom{n}{k}$ then the estimate for the dimension $m$ in* (ii) *is tight. In particular, the estimate for $m$ in* (i) *is tight in terms of $k, p$.*

Remark 1.3. The maps $H$ in Theorem 1.2(i), (ii) are random and we shall show in Section 2 that they satisfy the conclusions with high probability estimates, which is an important ingredient in applications.

*Proof.* The proofs of (i) and (ii) appear in Section 2.1. We now prove (iii). Recall that the Grassmann manifold $\mathcal{G}_{m,k}$ is the collection of all $k$-dimensional subspaces of $\mathbb{R}^m$ equipped with the metric

$$\rho(V,W) = \max_{v \in V \cap S^{m-1}} d(v, W \cap S^{m-1})$$

where $V, W$ are $k$-dimensional subspaces of $\mathbb{R}^m$ and the metric $d$ is Euclidean.

For $A, B \subset \mathbb{R}^n$ denote by $d_{\mathrm{aff}}(A,B)$ the distance between the affine subspaces spanned by $A$ and by $B$. If $A = \{x\}$ then we write $d_{\mathrm{aff}}(x,B)$.

Let $D \ge 2$ and $p \le \binom{n}{k}$. We may assume $k \le p$. Let $A = \{0, e_1, \dots, e_n\}$ where $\{e_i\}_{i=1}^n$ is the standard basis of $\ell_2^n$. Choose $p$ distinct subsets $B_1, \dots, B_p$ of $A$, each with $k+1$ elements one of which is zero. For every $1 \le i \ne j \le p$ choose $v_{i,j} \in B_i \setminus B_j$ and for every $1 \le i \le p$ choose an ordering $b_{i,0}, b_{i,1}, \dots, b_{i,k}$ of $B_i$ where $b_{i,0} = 0$. Note that $d_{\mathrm{aff}}(v_{i,j}, B_j) = 1$ for every $1 \le i \ne j \le p$.

First we prove a lower bound on the dimension for maps that distort affine distances by $D$ at most: Assume $F : A \to \ell_2^m$ is any map which satisfies $F(0) = 0$ and such that for all $1 \le i \ne j \le p$,

$$\text{(1)} \qquad \frac{1}{D} \le \frac{d_{\mathrm{aff}}(F(v_{i,j}), F(B_j))}{d_{\mathrm{aff}}(v_{i,j}, B_j)} = d_{\mathrm{aff}}(F(v_{i,j}), F(B_j)) \le 1,$$

$$\frac{1}{D} \le \frac{d_{\mathrm{aff}}(F(v_{i,j}), 0)}{d_{\mathrm{aff}}(v_{i,j}, 0)} = \frac{d(F(v_{i,j}), 0)}{d(v_{i,j}, 0)} = \|F(v_{i,j})\|_2 \le 1,$$

and for all $1 \le i \le p$ and $1 \le l \le k$,

$$\text{(2)} \qquad \frac{1}{D} \le \frac{d_{\mathrm{aff}}(F(b_{i,l}), F(\{b_{i,0}, \dots, b_{i,l-1}\}))}{d_{\mathrm{aff}}(b_{i,l}, \{b_{i,0}, \dots, b_{i,l-1}\})} \le 1.$$

Thus $F$ has distortion at most $D$ for at most $2p(p-1) + kp \le 3p^2$ affine distances. Put $V_i = \operatorname{span}(F(B_i))$. Then by (2), $V_i$ are $k$-dimensional subspaces of $\ell_2^m$. Consider $V_i$ as elements of the Grassmann manifold $\mathcal{G}_{m,k}$. Then by (1), $\{V_i\}_{i=1}^p$ satisfy $\rho(V_i, V_j) \ge 1/D$ for every $1 \le i \ne j \le p$. Hence the Grassmann manifold $\mathcal{G}_{m,k}$ contains $p$ disjoint balls of radius $1/2D$ and therefore $p \cdot \mu_{m,k}(B_\rho(V_1, 1/2D)) \le 1$ where $\mu_{m,k}$ is the normalized Haar measure on $\mathcal{G}_{m,k}$. It is known that there is a universal constant $C>0$ such that for every $0<r<1/2$,

$$\mu_{m,k}(B_\rho(W,r)) \geq Cr^{m-k}/\sqrt{m}.$$

Hence $m \geq c(k+\ln p/\ln D)$ for some universal constant $c>0$. We now observe that if $F:\ell_2^n \to \ell_2^m$ is linear and for all $1\leq i\neq j\leq p$ and $x,y$ in the affine span of $\{v_{i,j}\}\cup B_j$,

$$\frac{1}{D}\,\|x-y\|_2 \leq \|F(x)-F(y)\|_2 \leq \|x-y\|_2,$$

then $F:A\to \ell_2^m$ satisfies (1) and (2) and therefore $m\geq c(k+\ln p/\ln D)$. In particular, for $D=2$ we conclude that the estimate for $m$ in (i) is tight in terms of $k,p$. ∎

Remark 1.4. The proof shows in fact that if $p\leq \binom{n}{k}$ then there is a choice of an $(n+1)$-point set $A$ and $3p^2$ distinct pairs $\{(x_i,Y_i)\}_{i=1}^{3p^2}$ where $x_i\in A$ and $Y_i\subset A$ with $|Y_i|\leq k$ such that any map $F:A\to\ell_2^m$ which satisfies

$$\frac{1}{D} \leq \frac{d_{\mathrm{aff}}(F(x_i),F(Y_i))}{d_{\mathrm{aff}}(x_i,Y_i)} \leq 1$$

for every $1\leq i\leq 3p^2$ must satisfy $m\geq c(k+\ln p/\ln D)$.

Using Theorem 1.2 we obtain as a corollary a strengthened form of a theorem of Magen [18]:

Corollary 1.5. *Let $1\leq k\leq n/4$ and let $A$ be an $n$-point subset of $\ell_2^n$.*

(i) *Let $0<\varepsilon<1$. There is a linear map $H:\ell_2^n\to\ell_2^m$ with $m=O(\varepsilon^{-2}k\ln(n/k))$ such that for every $Y\subset A$ with at most $k$ elements and all $x,y$ in the affine subspace spanned by $Y$,*

$$(1-\varepsilon)\|x-y\|_2 \leq \|H(x)-H(y)\|_2 \leq (1+\varepsilon)\|x-y\|_2.$$

*Hence, for any two subsets $X,Y\subset A$ with $|X\cup Y|\leq k$,*

$$(1-\varepsilon)d_{\mathrm{aff}}(X,Y) \leq d_{\mathrm{aff}}(H(X),H(Y)) \leq (1+\varepsilon)d_{\mathrm{aff}}(X,Y).$$

(ii) *Let $D\geq 2$. There is a linear map $H:\ell_2^n\to\ell_2^m$ with $m=O\big(\frac{k\ln(n/k)}{\ln D}\big)$ such that for every $Y\subset A$ with at most $k$ elements and all $x,y$ in the affine subspace spanned by $Y$,*

$$\frac{1}{D}\,\|x-y\|_2 \leq \|H(x)-H(y)\|_2 \leq \|x-y\|_2.$$

*Hence, for any two subsets $X,Y\subset A$ with $|X\cup Y|\leq k$,*

$$\frac{1}{D}\,d_{\mathrm{aff}}(X,Y) \leq d_{\mathrm{aff}}(H(X),H(Y)) \leq d_{\mathrm{aff}}(X,Y).$$

(iii) *The estimate for the dimension $m$ in* (ii) *is tight. In particular, the estimate for $m$ in* (i) *is tight in terms of $k,p$.*

*Proof.* Put $p = \binom{n}{k}$. Let $\{W_l\}_{l=1}^p$ be the affine subspaces of $\mathbb{R}^n$ which are spanned by $k$ elements of $A$. Theorem 1.2, applied to the parallel subspaces of $\{W_l\}_{l=1}^p$, and the simple inequality $\frac{1}{2}k\ln\frac{n}{k} \le \ln\binom{n}{k} \le 3k\ln\frac{n}{k}$ conclude the proof. ∎

**2. Embedding Euclidean subsets into Banach spaces.** We begin with some preliminaries. Henceforth $g_{i,j}, g_j, h_i, g$ will denote independent $\mathcal{N}(0,1)$ random variables, i.e. independent standard Gaussian variables. For a positive integer $m$ define

$$a_m = \mathbb{E}\sqrt{\sum_{j=1}^m g_j^2} = \frac{\sqrt{2}\,\Gamma((m+1)/2)}{\Gamma(m/2)}.$$

It is well known that

$$\sqrt{\frac{m}{m+1}} \le \frac{a_m}{\sqrt{m}} \le 1\ . \tag{3}$$

We shall need estimates on the expectations of random variables which are functions of Gaussian variables and concentration estimates. The first result we need is the Maurey–Pisier concentration theorem (see [20, pp. 176–184]).

Theorem 2.1. *Let $F : \ell_2^m \to \mathbb{R}$ have Lipschitz constant $\sigma$ and let $\{g_j\}_{j=1}^m$ be independent standard Gaussian variables. Then for every $\lambda > 0$,*

$$\Pr(F(g_1,\dots,g_m) - \mathbb{E}F(g_1,\dots,g_m) < \lambda) \ge 1 - \exp\left(-\frac{\lambda^2}{2\sigma^2}\right), \tag{4}$$

$$\Pr(F(g_1,\dots,g_m) - \mathbb{E}F(g_1,\dots,g_m) > -\lambda) \ge 1 - \exp\left(-\frac{\lambda^2}{2\sigma^2}\right). \tag{5}$$

Let $u : \ell_2^m \to Y$ be a linear map from Euclidean space into a Banach space $Y$. The *$\ell$-norm* of $u$, as defined by Figiel and Tomczak-Jaegermann [9], is $\ell(u) = \mathbb{E}\|\sum_{j=1}^m g_j u(e_j)\|_Y$ where $\{e_j\}_{j=1}^m$ is the standard basis for $\ell_2^m$. The $\ell$-norm of $Y$ is

$$\ell(Y) = \sup\{\ell(u) \mid u : \ell_2^m \to Y,\ \|u\| = 1 \text{ and } m \ge 1\}.$$

If $Y$ is finite-dimensional then the supremum is a maximum and it is attained for some $u : \ell_2^m \to Y$ where $m = \dim(Y)$. It is known that $\ell(\ell_2^m) = a_m$, $\ell(\ell_p^m) \approx \sqrt{p}\,m^{1/p}$ whenever $2 < p \le \ln m$, and $\ell(\ell_p^m) \approx \sqrt{m}$ whenever $1 \le p < 2$. For a general $m$-dimensional Banach space $Y$ we have $c_1\sqrt{\ln m} \le \ell(Y) \le a_m$. Given Banach spaces $X, Y$ and $x^* \in X^*$ and $y \in Y$, the linear operator $x^* \otimes y : X \to Y$ is defined by $(x^* \otimes y)(x) = \langle x, x^*\rangle y$ where $\langle x, x^*\rangle = x^*(x)$.

The second result we need is a restatement of results by Gordon [11, 13].

Theorem 2.2. *Let $X, Y$ be Banach spaces, let $S \subset X$ be a compact subset and $\{x_i^*\}_{i=1}^n \subset X^*$ and $\{y_j\}_{j=1}^m \subset Y$. Define linear maps $v : X \to \ell_2^n$ by*

*$v(x)=\{\langle x,x_i^*\rangle\}_{i=1}^n$ and $u:\ell_2^m\to Y$ by $u(e_j)=y_j$. Assume $g_{i,j},g_j,h_i,g$ are independent standard Gaussian variables. Put $G=\sum_{i=1}^n\sum_{j=1}^m g_{i,j}\,x_i^*\otimes y_j$. Then*

$$\ell(u)\min_{x\in S}\|v(x)\|_2-\|u\|\,\mathbb{E}\max_{x\in S}\sum_{i=1}^n\langle x,x_i^*\rangle h_i\le \mathbb{E}\min_{x\in S}\|G(x)\|_Y$$

$$\le \mathbb{E}\max_{x\in S}\|G(x)\|_Y\le \ell(u)\max_{x\in S}\|v(x)\|_2+\|u\|\,\mathbb{E}\max_{x\in S}\sum_{i=1}^n\langle x,x_i^*\rangle h_i.$$

We will use Theorem 2.2 and (7) in the case where $X=\ell_2^n$ and $v$ is the identity and $S\subset S^{n-1}$ is closed. For a closed subset $S\subset S^{n-1}$ define $\mathbb{E}(S)=\mathbb{E}\max_{x\in S}\sum_{i=1}^n\langle x,e_i\rangle h_i$. Then

$$\text{(6)}\qquad \ell(u)-\|u\|\,\mathbb{E}(S)\le \mathbb{E}\min_{x\in S}\|G(x)\|_Y \le \mathbb{E}\max_{x\in S}\|G(x)\|_Y\le \ell(u)+\|u\|\,\mathbb{E}(S).$$

It is also well known that in this case

$$\text{(7)}\qquad \max(\ell(u),\|u\|\mathbb{E}(S))\le \mathbb{E}\max_{x\in S}\|G(x)\|_2.$$

Let $Y$ be a Banach space. Following the definition in [14] we say that $\{y_j\}_{j=1}^m\subset Y$ satisfies a $(C,s)$-*estimate* for $C>0$ and $s>0$ if for all $\{a_j\}_{j=1}^m\in\mathbb{R}^m$ and $h\le m$,

$$Ch^{-1/s}\Big(\sum_{j=1}^h (a_j^*)^2\Big)^{1/2}\le\Big\|\sum_{j=1}^m a_jy_j\Big\|\le\Big(\sum_{j=1}^m a_j^2\Big)^{1/2}$$

where $\{a_j^*\}_{j=1}^m$ is a decreasing rearrangement of $\{|a_j|\}_{j=1}^m$. By a result of Bourgain and Szarek [5], there exists $C>0$ such that any $n$-dimensional Banach space contains a sequence $u_1,\dots,u_N$, with $N\ge n/2$, which satisfies a $(C,2)$-estimate.

In Section 2.1 we shall use the following theorem which generalizes results by Gordon [13] and by Gordon, Guédon, Meyer and Pajor [14] who needed it for special cases.

THEOREM 2.3. *There are constants $c_1,c_2,d_1,d_2$ such that given a closed $S\subset S^{n-1}$ the following holds:*

(i) *Assume $Y$ is an $m$-dimensional Banach space for which $\varepsilon^{-1}\mathbb{E}(S)\le\ell(Y)$ for some $0<\varepsilon\le 1/2$. Choose $u:\ell_2^m\to Y$ for which $\ell(u)=\ell(Y)$ and $\|u\|=1$. Put $y_j=u(e_j)$ where $\{e_j\}_{j=1}^m$ is the standard basis for $\ell_2^m$ and let*

$$H=\frac{1}{\ell(u)}G=\frac{1}{\ell(u)}\sum_{i=1}^n\sum_{j=1}^m g_{i,j}\,e_i\otimes y_j:\ell_2^n\to Y.$$

*Then the probability that*

$$(1-\varepsilon)^2 \le \|H(x)\|_Y \le (1+\varepsilon)^2$$

*for every* $x \in S$ *is at least* $1 - 2\exp(-\varepsilon^2(1-\varepsilon)^2\ell(Y)^2/2)$.

(ii) *Assume* $Y$ *is a Banach space and* $\{y_j\}_{j=1}^m \subset Y$ *satisfies a* $(C,s)$*-estimate for some* $C > 0$ *and* $s \ge 2$ *and put* $1/q = \max(1/2 - 1/s, 1/\ln m)$. *Let*

$$G = \sum_{i=1}^n \sum_{j=1}^m g_{i,j}\, e_i \otimes y_j : \ell_2^n \to Y.$$

(a) *If* $\frac{1}{2}\ell(Y) \le \mathbb{E}(S) \le c_1 C\sqrt{qm}\, e^{-q/2}$ *then with positive probability,*

$$\frac{\max_{x\in S}\|G(x)\|_Y}{\min_{x\in S}\|G(x)\|_Y} \le \frac{d_1\mathbb{E}(S)}{C\sqrt{q}\, m^{1/q}}.$$

(b) *If* $c_1 C\sqrt{qm}\, e^{-q/2} \le \mathbb{E}(S) \le c_2\sqrt{m}$ *then with positive probability,*

$$\frac{\max_{x\in S}\|G(x)\|_Y}{\min_{x\in S}\|G(x)\|_Y} \le \frac{d_2\mathbb{E}(S)^{1-2/q}}{C(\ln(m/\mathbb{E}(S)^2))^{1/2}}.$$

*Proof.* (i) Using (6) and $\|u\| = 1$ and $\ell(u) = \ell(Y)$ and $\mathbb{E}(S) \le \varepsilon\ell(Y)$ we obtain

$$\ell(Y)(1-\varepsilon) \le \mathbb{E}\min_{x\in S}\|G(x)\|_Y \le \mathbb{E}\max_{x\in S}\|G(x)\|_Y \le \ell(Y)(1+\varepsilon). \tag{8}$$

For fixed $x = \{\langle x, e_i\rangle\}_{i=1}^n \in S^{n-1}$ the function $\|\sum_{i=1}^n \sum_{j=1}^m t_{i,j}\langle x, e_i\rangle y_j\|_Y$ from $\ell_2^{n\times m}$ to $\mathbb{R}$ has Lipschitz constant at most one. Indeed,

$$\begin{aligned}
&\Big|\Big\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x, e_i\rangle y_j\Big\|_Y - \Big\|\sum_{i=1}^n\sum_{j=1}^m s_{i,j}\langle x, e_i\rangle y_j\Big\|_Y\Big| \\
&\quad \le \Big\|\sum_{i=1}^n\sum_{j=1}^m (t_{i,j}-s_{i,j})\langle x, e_i\rangle y_j\Big\|_Y \\
&\quad = \sup_{\|y^*\|=1}\Big\langle \sum_{i=1}^n\sum_{j=1}^m (t_{i,j}-s_{i,j})\langle x, e_i\rangle y_j, y^*\Big\rangle \\
&\quad = \sup_{\|y^*\|=1}\sum_{i=1}^n\sum_{j=1}^m (t_{i,j}-s_{i,j})\langle x, e_i\rangle\langle y_j, y^*\rangle \\
&\quad \le \sup_{\|y^*\|=1}\Big(\sum_{i=1}^n\sum_{j=1}^m (t_{i,j}-s_{i,j})^2\Big)^{1/2}\Big(\sum_{i=1}^n\sum_{j=1}^m \langle x, e_i\rangle^2\langle y_j, y^*\rangle^2\Big)^{1/2} \\
&\quad = \|\{t_{i,j}\}_{i,j} - \{s_{i,j}\}_{i,j}\|_2\|u\| = \|\{t_{i,j}\}_{i,j} - \{s_{i,j}\}_{i,j}\|_2.
\end{aligned}$$

This clearly implies that $\max_{x\in S}\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x,e_i\rangle y_j\|_Y$ has Lipschitz constant at most one. We now show that $\min_{x\in S}\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x,e_i\rangle y_j\|_Y$ also has Lipschitz constant one:

$$\Big\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x,e_i\rangle y_j\Big\|_Y \le \Big\|\sum_{i=1}^n\sum_{j=1}^m s_{i,j}\langle x,e_i\rangle y_j\Big\|_Y + \|\{t_{i,j}\}_{i,j}-\{s_{i,j}\}_{i,j}\|_2,$$

hence

$$\min_{x\in S}\Big\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x,e_i\rangle y_j\Big\|_Y \le \min_{x\in S}\Big\|\sum_{i=1}^n\sum_{j=1}^m s_{i,j}\langle x,e_i\rangle y_j\Big\|_Y + \|\{t_{i,j}\}_{i,j}-\{s_{i,j}\}_{i,j}\|_2,$$

which proves the claim.

We now apply Theorem 2.1 to $\max_{x\in S}\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x,e_i\rangle y_j\|_Y$ and $\min_{x\in S}\|\sum_{i=1}^n\sum_{j=1}^m t_{i,j}\langle x,e_i\rangle y_j\|_Y$, and using (8) we obtain

$$\Pr(\max_{x\in S}\|G(x)\|_Y < (1+\varepsilon)^2\ell(Y)) \ge 1-\exp(-\varepsilon^2(1-\varepsilon)^2\ell(Y)^2/2),$$
$$\Pr(\min_{x\in S}\|G(x)\|_Y > (1-\varepsilon)^2\ell(Y)) \ge 1-\exp(-\varepsilon^2(1-\varepsilon)^2\ell(Y)^2/2),$$

which concludes the proof of (i).

(ii) Note that if $1/2-1/s<1/\ln m$ then $\{y_j\}_{j=1}^m$ satisfies a $(C/e,s')$-estimate where $1/s'=1/2-1/\ln m$. Hence we may assume that $1/q=1/2-1/s\ge 1/\ln m$, i.e. $q\le\ln m$.

Define $u:\ell_2^m\to Y$ by $u(e_j)=y_j$. Then $\|u\|\le 1$ because $\{y_j\}_{j=1}^m$ satisfies a $(C,s)$-estimate.

Fix $1\le h\le m$ to be determined later. Define a new norm on $\operatorname{span}\{y_j\}_{j=1}^m$ by

$$\Big\|\sum_{j=1}^m a_j y_j\Big\|_{(h)} = Ch^{-1/s}\Big(\sum_{j=1}^h (a_j^*)^2\Big)^{1/2}.$$

Then $\|G(x)\|_Y \ge \|G(x)\|_{(h)}$ for every $x\in S$ because $\{y_j\}_{j=1}^m$ satisfies a $(C,s)$-estimate and therefore

$$\mathbb{E}\min_{x\in S}\|G(x)\|_Y \ge \mathbb{E}\min_{x\in S}\|G(x)\|_{(h)}.$$

Applying (6) to $G:\ell_2^n\to(\operatorname{span}\{y_j\}_{j=1}^m,\|\cdot\|_{(h)})$ we obtain

$$\begin{aligned}\mathbb{E}\min_{x\in S}\|G(x)\|_{(h)} &\ge \mathbb{E}\Big\|\sum_{j=1}^m g_j y_j\Big\|_{(h)} - Ch^{-1/s}\mathbb{E}(S)\\ &\ge Ch^{-1/s}\Big(\mathbb{E}\Big(\sum_{j=1}^h (g_j^*)^2\Big)^{1/2} - \mathbb{E}(S)\Big)\\ &\ge Ch^{-1/s}\Big(c_0h^{1/2}\Big(\ln\Big(1+\frac{m}{h}\Big)\Big)^{1/2} - \mathbb{E}(S)\Big),\end{aligned}$$

where the last inequality is a classical estimate of $\mathbb{E}(\sum_{j=1}^{h}(g_j^*)^2)^{1/2}$ (see [15]). Applying (6) to $G : \ell_2^n \to Y$ we obtain

$$\mathbb{E}\max_{x\in S}\|G(x)\|_Y \le \ell(u) + \mathbb{E}(S) \le \ell(Y) + \mathbb{E}(S) \le 3\mathbb{E}(S).$$

Hence

$$\frac{\mathbb{E}\max_{x\in S}\|G(x)\|_Y}{\mathbb{E}\min_{x\in S}\|G(x)\|_Y} \le \frac{3\mathbb{E}(S)}{Ch^{-1/s}(c_0 h^{1/2}(\ln(1+m/h))^{1/2} - \mathbb{E}(S))}.$$

To prove (a) choose $h \approx me^{-q}$. Then

$$\frac{\mathbb{E}\max_{x\in S}\|G(x)\|_Y}{\mathbb{E}\min_{x\in S}\|G(x)\|_Y} \le \frac{d_1\mathbb{E}(S)}{C\sqrt{q}\, m^{1/q}}.$$

To prove (b) choose $h \approx \mathbb{E}(S)^2$. Then

$$\frac{\mathbb{E}\max_{x\in S}\|G(x)\|_Y}{\mathbb{E}\min_{x\in S}\|G(x)\|_Y} \le \frac{d_2\mathbb{E}(S)^{2/s}}{C(\ln(m/\mathbb{E}(S)^2))^{1/2}}. \blacksquare$$

REMARK 2.4. A closer look at the proof of (ii) shows that the probability of the conclusion of (a) is at least $1 - \exp(-c\, C^2 q\, m^{2/q})$ and the probability of the conclusion of (b) is at least $1 - \exp(-c\, C^{2+2/q} q\, m^{1/q})$.

**2.1.** *On neighborhoods of affine subspaces.* We now consider particular subsets of $S^{n-1}$. We shall need the following proposition in order to estimate various distortions in various cases associated with embeddings of affine subspaces into general Banach spaces.

PROPOSITION 2.5. *Let $1 \le k \le n$ and let $\{W_l\}_{l=1}^p$ be linear subspaces of dimension at most $k$ in $\ell_2^n$. Fix $0 \le r < 1$. Put*

$$S = \{x \in S^{n-1} : \textit{there exists } 1 \le l \le p \textit{ such that } d(x, W_l \cap S^{n-1}) \le r\}.$$

*Then*

$$\mathbb{E}(S) = \mathbb{E}\max_{x\in S}\sum_{i=1}^{n}\langle x, e_i\rangle h_i \le 3(\sqrt{\ln p} + a_k + r(\sqrt{\ln p} + a_{n-k})).$$

*Proof.* Fix $1 \le l \le p$. Put

$$S_l = \{x \in S^{n-1} : d(x, W_l \cap S^{n-1}) \le r\}.$$

Choose $U_l \in O(n)$ such that $U_l(W_l) = \mathbb{R}^k \subset \mathbb{R}^n$. Then

$$U_l(S_l) = \{x \in S^{n-1} : d(x, S^{k-1}) \le r\}.$$

Let $Z_{l,i}$ be the entries of $U_l(h)$ where $h = (h_i)_{i=1}^n$ and $h_i$ are independent standard Gaussian variables. Then $\{Z_{l,i}\}_{i=1}^n$ are independent standard Gaussian variables. We have

$$
\begin{aligned}
(9) \qquad \mathbb{E}(S) &= \mathbb{E}\max_{x\in S}\sum_{i=1}^{n}\langle x, e_i\rangle h_i = \mathbb{E}\max_{1\le l\le p}\max_{x\in S_l}\langle x, h\rangle \\
&= \mathbb{E}\max_{1\le l\le p}\max_{x\in S_l}\langle U_l(x), U_l(h)\rangle = \mathbb{E}\max_{1\le l\le p}\max_{x\in U_l(S_l)}\langle x, U_l(h)\rangle \\
&\le \mathbb{E}\max_{1\le l\le p}\Bigg(\sqrt{\sum_{i=1}^{k} Z_{l,i}^2} + r\sqrt{\sum_{i=k+1}^{n} Z_{l,i}^2}\Bigg) \\
&\le \mathbb{E}\max_{1\le l\le p}\sqrt{\sum_{i=1}^{k} Z_{l,i}^2} + r\mathbb{E}\max_{1\le l\le p}\sqrt{\sum_{i=k+1}^{n} Z_{l,i}^2}\ .
\end{aligned}
$$

Let $X_l = \sqrt{\sum_{i=1}^{k} Z_{l,i}^2}$ . Using (5) we find that for all $1 \le l \le p$ and $t > a_k$,

$$
\begin{aligned}
(10) \qquad \Pr(X_l \ge t) &= \Pr\Big(X_l \ge \frac{t}{a_k}\, a_k\Big) \\
&\le \exp\bigg(-\frac{1}{2}\Big(\frac{t}{a_k} - 1\Big)^2 a_k^2\bigg) = \exp(-(t-a_k)^2/2).
\end{aligned}
$$

Using (10) we infer that for $a > a_k$,

$$
\begin{aligned}
\mathbb{E}\max_{1\le l\le p} X_l &= \int_0^\infty \Pr(\max_{1\le l\le p} X_l \ge t)\, dt \\
&= \int_0^a \Pr\Big(\bigcup_{1\le l\le p}(X_l \ge t)\Big)\, dt + \int_a^\infty \Pr\Big(\bigcup_{1\le l\le p}(X_l \ge t)\Big) dt \\
&\le a + p\int_a^\infty \Pr(X_1 \ge t)\, dt \le a + p\int_a^\infty \exp(-(t-a_k)^2/2)\, dt \\
&\le a + p\int_a^\infty \frac{t-a_k}{a-a_k}\exp(-(t-a_k)^2/2)\, dt \\
&= a + \frac{p}{a-a_k}\exp(-(a-a_k)^2/2).
\end{aligned}
$$

Choose $a = 2(\sqrt{\ln p} + a_k)$ to obtain

$$
(11) \qquad \mathbb{E}\max_{1\le l\le p}\sqrt{\sum_{i=1}^{k} Z_{l,i}^2} = \mathbb{E}\max_{1\le l\le p} X_l \le 3(\sqrt{\ln p} + a_k).
$$

Similarly

$$
(12) \qquad \mathbb{E}\max_{1\le l\le p}\sqrt{\sum_{i=k+1}^{n} Z_{l,i}^2} \le 3(\sqrt{\ln p} + a_{n-k}).
$$

Combining (9), (11) and (12) we obtain

$$\mathbb{E}(S) = \mathbb{E}\max_{x\in S}\sum_{i=1}^{n}\langle x, e_i\rangle h_i \le 3\big(\sqrt{\ln p} + a_k + r(\sqrt{\ln p} + a_{n-k})\big). \ \blacksquare$$

Combining Theorem 2.3(i) and Proposition 2.5 yields

THEOREM 2.6. *Fix $0 < \varepsilon < 1/2$ and $1 \le k \le n$. Let $\{W_l\}_{l=1}^{p}$ be linear subspaces of dimension at most $k$ in $\ell_2^n$ and let $Y$ be a Banach space which satisfies $\ell(Y) \ge 3\varepsilon^{-1}(a_k + \sqrt{\ln p})$. Then there is a linear map $H : \ell_2^n \to Y$ such that for all $1 \le l \le p$ and $x, y \in W_l$,*

$$(1-\varepsilon)\|x-y\|_2 \le \|H(x) - H(y)\|_Y \le (1+\varepsilon)\|x-y\|_2.$$

REMARK 2.7. Since $\ell(\ell_2^m) = a_m \sim \sqrt{m}$ we obtain Theorem 1.2(i). Moreover, Proposition 2.5 actually implies that Theorem 2.6 holds for every affine subspace $W$ of $\ell_2^n$ whose parallel subspace $Y$ through the origin is within distance $O((\sqrt{k} + \sqrt{\ln p})/\sqrt{n})$ to one of the subspaces $\{W_l\}_{l=1}^{p}$, or equivalently, Theorem 2.6 holds for every affine subspace $W$ of $\ell_2^n$ whose parallel subspace $Y$ through the origin is in $\bigcup_{l=1}^{p} B_\rho(W_l, O((\sqrt{k} + \sqrt{\ln p})/\sqrt{n}))$ where $\rho$ is the metric on the Grassmann manifold.

The result below is a quantitative version of Theorem 1.2(ii) for large distortions $D$, and by Theorem 1.2(iii), the estimate on $m$ is tight. For $m, n$ let $G = \sum_{i=1}^{n}\sum_{j=1}^{m} g_{i,j}\, e_i \otimes e_j : \ell_2^n \to \ell_2^m$ where $g_{i,j}$ are independent standard Gaussian variables.

THEOREM 2.8. *There is a positive constant $c$ such that the following holds: Given $D \ge c$ and $1 \le k \le n$ and $p$ subspaces $\{W_l\}_{l=1}^{p}$ of dimension at most $k$ in $\ell_2^n$ and any $m \ge 5(k + \ln p/\ln D)$, there is a number $E$ such that the probability that for all $1 \le l \le p$ and $x, y \in W_l$,*

$$E\,\frac{\sqrt{\ln D}}{D}\,\|x-y\|_2 \le \|G(x) - G(y)\|_2 \le 4E\sqrt{\ln D}\,\|x-y\|_2$$

*is at least $1 - 2/p$.*

*Proof.* Put $S = \bigcup_{l=1}^{p}(W_l \cap S^{n-1})$ and $E = \mathbb{E}\max_{x\in S}\|G(x)\|_2$. We shall prove that the probability that

$$E\,\frac{\sqrt{\ln D}}{D} \le \min_{x\in S}\|G(x)\|_2 \le \max_{x\in S}\|G(x)\|_2 \le 4E\sqrt{\ln D}$$

is at least $1 - 2/p$, which will prove the theorem.

Since in $S^{k-1}$, for every $0 < \varepsilon < 1$ there is an $\varepsilon$-net $N$ with $|N| \le (3/\varepsilon)^k$, it follows that there is a $1/4D$-net $\{x_h\}_{h\in A}$ of $S$ with $|A| \le p(12D)^k$.

Put $\Lambda = \big\{\max_{x\in S}\|G(x)\|_2 \le 4E\sqrt{\ln D}\big\}$. Using conditional probability we obtain

$$\text{(13)}\qquad \Pr\Big(E\,\frac{\sqrt{\ln D}}{D}\le \min_{x\in S}\|G(x)\|_2\le \max_{x\in S}\|G(x)\|_2\le 4E\sqrt{\ln D}\Big)$$
$$=\Pr\Big(E\,\frac{\sqrt{\ln D}}{D}\le \min_{x\in S}\|G(x)\|_2\ \Big|\ \Lambda\Big)\cdot \Pr(\Lambda)$$
$$\ge \Pr\Big(\bigcap_{h\in A}\Big\{2E\,\frac{\sqrt{\ln D}}{D}\le \|G(x_h)\|_2\Big\}\ \Big|\ \Lambda\Big)\Pr(\Lambda)$$
$$=\Pr\Big(\bigcap_{h\in A}\Big\{2E\,\frac{\sqrt{\ln D}}{D}\le \|G(x_h)\|_2\Big\}\cap \Lambda\Big)$$
$$\ge 1-\Pr\Big(\bigcup_{h\in A}\Big\{\|G(x_h)\|_2\le 2E\frac{\sqrt{\ln D}}{D}\Big\}\Big)-\Pr(\Lambda^c).$$

Using (6), (7) with $u$ the identity on $\ell_2^m$ and then $\ell(u)=a_m$, and $\ln p\le \frac{1}{5}m\ln D$ and Proposition 2.5, we obtain

$$\text{(14)}\qquad \frac{1}{2}\sqrt{m}\le a_m\le E\le a_m+\mathbb{E}(S)\le a_m+3a_k+3\sqrt{\ln p}\le 6\sqrt{m\ln D}$$

and therefore as $\ln p\le \frac{1}{5}m\ln D$, by (14) and Theorem 2.1,

$$\text{(15)}\qquad \Pr(\Lambda^c)\le \exp(-(4\sqrt{\ln D}-1)^2a_m^2/2)\le \exp(-\ln p)=1/p.$$

We now estimate $\Pr(\bigcup_{h\in A}\{\|G(x_h)\|_2\le 2c_1E\sqrt{\ln D}/D\})$. We shall need the following estimate obtained using the formula $\mathrm{Vol}_m(B_2^m)=\pi^{m/2}/\Gamma(1+m/2)$ and the well known estimate $\Gamma(1+t)\ge (t/e)^t$: For all $x\in S^{n-1}$ and $\lambda>0$,

$$\text{(16)}\qquad \Pr(\|G(x)\|_2\le\lambda)=\frac{1}{(2\pi)^{m/2}}\int_{\lambda B_2^m}\exp\Big(-\frac{1}{2}\|y\|_2^2\Big)\,dy$$
$$\le \frac{\mathrm{Vol}_m(\lambda B_2^m)}{(2\pi)^{m/2}}=\frac{\lambda^m\,\mathrm{Vol}_m(B_2^m)}{(2\pi)^{m/2}}\le \exp\Big(m\ln\lambda-\frac{m}{2}\ln\frac{m}{e}\Big).$$

Using (14), (16) and $k\le m/5$ and a large enough $c$ we obtain

$$\text{(17)}\qquad \Pr\Big(\bigcup_{h\in A}\Big\{\|G(x_h)\|_2\le 2E\,\frac{\sqrt{\ln D}}{D}\Big\}\Big)$$
$$\le p(12D)^k\cdot\exp\Big(m\ln\Big(\frac{2E\sqrt{\ln D}}{D}\Big)-\frac{m}{2}\ln\frac{m}{e}\Big)$$
$$\le \exp\Big(m\ln\Big(\frac{12\sqrt{m}\ln D}{D}\Big)-\frac{m}{2}\ln\frac{m}{e}+\ln p+k\ln(12D)\Big)$$
$$\le \exp\Big(-\frac{1}{5}\,m\ln D\Big)\le \exp(-\ln p)=\frac{1}{p},$$

and substituting (15) and (17) into (13) concludes the proof. ∎

Combining Theorems 2.6 and 2.8 gives the following corollary.

Corollary 2.9. *There are positive constants $c_1, c_2$ such that given $D \geq c_1$ and $1 \leq k \leq n$ and $p$ affine subspaces $\{W_l\}_{l=1}^p$ of dimension at most $k$ in $\ell_2^n$ and a Banach space $Y$ for which $\ell(Y)^2 \geq c_2(k + \ln p/\ln D)$, there is a map $H : \ell_2^n \to Y$ such that for all $1 \leq l \leq p$ and $x, y \in W_l$,*

$$\frac{1}{D}\|x-y\|_2 \leq \|H(x) - H(y)\|_Y \leq \|x-y\|_2.$$

**3. Random mappings into Euclidean spaces.** In the special case when the target space is Euclidean, the question arises whether random variables other than Gaussian random variables can be used. This is a consequence of Proposition 2.5 above, a result of Mendelson, Pajor and Tomczak-Jaegermann [19] and a result of Talagrand [21]. We recall some definitions.

The $\psi_2$ *norm* of a random variable $X$ is

$$\|X\|_{\psi_2} = \inf\{C > 0 : \mathbb{E}\exp(X^2/C^2) \leq 2\}.$$

Definition 3.1. A probability measure $\mu$ on $\mathbb{R}^n$ is called *isotropic* if for every $y \in \mathbb{R}^n$ we have $\mathbb{E}|\langle X, y\rangle|^2 = \|y\|_2^2$ where $X$ is distributed according to $\mu$. We say that $\mu$ is a $\psi_2$ *probability measure* with constant $\alpha$ if additionally, for every $y \in \mathbb{R}^n$ we have $\|\langle X, y\rangle\|_{\psi_2} \leq \alpha\|y\|_2$.

Given a compact metric space $(T, d)$, define

$$\gamma_2(T) = \inf \sup_{t\in T} \sum_{s=0}^{\infty} 2^{s/2} d(t, T_s)$$

where the infimum is taken over all subsets $T_s \subset T$ with cardinality $|T_s| \leq 2^{2^s}$ and $|T_0| = 1$. The result of Talagrand [21] states that there are positive constants $c_1, c_2$ such that given a Gaussian process $\{X_t : t \in T\}$ for which $d(s,t)^2 = \mathbb{E}(X_t - X_s)^2$, we have

$$c_1\gamma_2(T) \leq \mathbb{E}\sup_{t\in T} X_t \leq c_2\gamma_2(T). \tag{18}$$

The results from [19] (Theorem 2.3 combined with Corollary 2.7) state:

Theorem 3.2. *There are positive constants $c_1, c_2$ for which the following holds: Let $T \subset S^{n-1}$ and $0 < \varepsilon < 1$, and let $\mu$ be an isotropic $\psi_2$ probability measure with constant $\alpha \geq 1$. For $m \geq 1$, let $X_1, \ldots, X_m$ be independent, distributed according to $\mu$, and define $\Gamma : \mathbb{R}^n \to \mathbb{R}^m$ to be the matrix whose rows are $X_1, \ldots, X_m$. If $m \geq c_1(\alpha^4/\varepsilon^2)\mathbb{E}(T)^2$ then with probability at least $1 - \exp(-c_2\varepsilon^2 m/\alpha^4)$, for every $x \in T$,*

$$1 - \varepsilon \leq \frac{1}{m}\|\Gamma x\|_2^2 \leq 1 + \varepsilon.$$

It is known that $n$-dimensional random vectors distributed uniformly on $S^{n-1}$ satisfy the conditions of Theorem 3.2 for some $\alpha$. It is also known that if $X$ is a random variable satisfying $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$ and $\|X\|_{\psi_2} \le \beta$ then the rows of the $m \times n$ random matrix whose entries are $X_{i,j}$ where $X_{i,j} \sim X$ are independent, satisfy the conditions of Theorem 3.2 for some $\alpha$ depending only on $\beta$. In particular, Bernoulli $\{-1, 1\}$ random variables and standard Gaussian random variables satisfy these conditions. Examples of random vectors whose coordinates are not independent identically distributed and satisfy the conditions of Theorem 3.2 are random vectors with the uniform distribution on the unit ball of $\ell_p^n$ for $p \ge 2$ (see [4]).

Combining Proposition 2.5 and Theorem 3.2 we obtain the following analogue of Theorem 2.3(i) when the target space is Euclidean.

THEOREM 3.3. *There are positive constants $c_1, c_2$ such that the following holds: Fix $0 < \varepsilon < 1$ and $1 \le k \le n$. Let $\{W_l\}_{l=1}^p$ be affine subspaces of dimension at most $k$ in $\ell_2^n$. Let $\mu$ be an isotropic $\psi_2$ probability measure with constant $\alpha \ge 1$ and $\Gamma_1, \dots, \Gamma_m$ be independent distributed according to $\mu$. Choose $m$ such that $m \ge c_1\alpha^4(k + \ln p)/\varepsilon^2$. Then the random matrix whose rows are $\Gamma_i$ is a linear map $\Gamma : \ell_2^n \to \ell_2^m$ such that with probability at least $1 - p^{-1/\alpha^4}$, for all $1 \le l \le p$ and $x, y \in W_l$,*

$$(1-\varepsilon)\|x-y\|_2 \le \|\Gamma(x) - \Gamma(y)\|_2 \le (1+\varepsilon)\|x-y\|_2.$$

*Proof.* Put $T = (\bigcup_{l=1}^n W_l) \cap S^{n-1}$. Using Proposition 2.5 we obtain $\mathbb{E}(T) \le 3(a_k + \sqrt{\ln p})$. Apply this estimate in Theorem 3.2. ∎

**Acknowledgements.** Research of Y. Gordon was partially supported by the France-Israel Scientific Exchange Program No. 3-4301, and by the Fund for the Promotion of Research at the Technion.

## References

[1] D. Achlioptas, *Database-friendly random projections: Johnson–Lindenstrauss with binary coins*, Special issue on PODS 2001 (Santa Barbara, CA), J. Comput. System Sci. 66 (2003), 671–687.

[2] N. Alon, *Problems and results in extremal combinatorics*, *Part 1*, unpublished manuscript.

[3] R. I. Arriaga and S. Vempala, *An algorithmic theory of learning*: *robust concepts and random projection*, in: 40th Annual Symposium on Foundations of Computer Science (New York, 1999), 1999, 616–623.

[4] F. Barthe and A. Koldobsky, *Extremal slabs in the cube and the Laplace transform*, Adv. Math. 174 (2003), 89–114.

[5] J. Bourgain and S. J. Szarek, *The Banach–Mazur distance to the cube and the Dvoretzky–Rogers factorization*, Israel J. Math. 62 (1988), 169–180.

[6] E. J. Candes and T. Tao, *Near-optimal signal recovery from random projections: universal encoding strategies*, IEEE Trans. Inform. Theory 52 (2006), 5406–5425.

[7] S. Dasgupta and A. Gupta, *An elementary proof of a theorem of Johnson and Lindenstrauss*, Random Structures Algorithms 22 (2003), 60–65.

[8] D. L. Donoho, *Compressed sensing*, IEEE Trans. Inform. Theory 52 (2006), 1289–1306.

[9] T. Figiel and N. Tomczak-Jaegermann, *Projections onto Hilbertian subspaces of Banach spaces*, Israel J. Math. 33 (1979), 155–171.

[10] P. Frankl and H. Maehara, *The Johnson–Lindenstrauss lemma and the sphericity of some graphs*, J. Combin. Theory Ser. B 44 (1988), 355–362.

[11] Y. Gordon, *Some inequalities for Gaussian processes and applications*, Israel J. Math. 50 (1985), 265–289.

[12] —, *Elliptically contoured distributions*, Probab. Theory Related Fields 76 (1987), 429–438.

[13] —, *On Milman's inequality and random subspaces which escape through a mesh in* $\mathbf{R}^n$, in: Geometric Aspects of Functional Analysis (1986/87), Lecture Notes in Math. 1317, Springer, 1988, 84–106.

[14] Y. Gordon, O. Guédon, M. Meyer and A. Pajor, *Random Euclidean sections of some classical Banach spaces*, Math. Scand. 91 (2002), 247–268.

[15] Y. Gordon, A. Litvak, C. Schütt and E. Werner, *On the minimum of several random variables*, Proc. Amer. Math. Soc. 134 (2006), 3665–3675.

[16] P. Indyk and R. Motwani, *Approximate nearest neighbors: towards removing the curse of dimensionality*, in: STOC '98 (Dallas, TX), 1999, 604–613.

[17] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, in: Conference in Modern Analysis and Probability (New Haven, CN, 1982), Contemp. Math. 26, Amer. Math. Soc., 1984, 189–206.

[18] A. Magen, *Dimensionality reductions that preserve volumes and distance to affine spaces, and their algorithmic applications*, in: Randomization and Approximation Techniques in Computer Science, Lecture Notes in Comput. Sci. 2483, Springer, 2002, 239–253.

[19] S. Mendelson, A. Pajor and N. Tomczak-Jaegermann, *Reconstruction and subgaussian operators in asymptotic geometric analysis*, Geom. Funct. Anal. 17 (2007), 1248–1282.

[20] G. Pisier, *Probabilistic methods in the geometry of Banach spaces*, in: Probability and Analysis (Varenna, 1985), Lecture Notes in Math. 1206, Springer, 1986, 167–241.

[21] M. Talagrand, *Majorizing measures without measures*, Ann. Probab. 29 (2001), 411–417.



Department of Mathematics
Technion
Haifa 32000, Israel
E-mail: alond@techunix.technion.ac.il
gordon@techunix.technion.ac.il



*Received December 2, 2008*
*Revised version April 10, 2009* (6477)