Two observations regarding embedding subsets of Euclidean spaces in normed spaces

Gideon Schechtman*

Abstract

This paper contains two results concerning linear embeddings of subsets of Euclidean space in low dimensional normed spaces. The first is an improvement of the known dependence on ε in Dvoretzky's theorem from order of ε^2 to order of ε (except for log factors). The second is a joint generalization of (Milman's version of) Dvoretzky's theorem and (a recent generalization by Klartag and Mendelson of) the Johnson-Lindenstrauss Lemma.

1 Introduction

Given a normed space X let

$$E(X) = \sup\{\mathbb{E}\Big(\|\sum_{i=1}^{n} g_i u(e_i)\|_X\Big) \ ; \ n \in \mathbb{N}, \ u : \ell_2^n \to X, \|u\| = 1\}. \tag{1}$$

Here and elsewhere in this paper $g_1, g_2, \ldots$ denote a sequence of independent standard Gaussian variables. The quantity $E(X)$ is, in Banach space terms, the ℓ norm of the identity on X. However, his fact will not be used here.

Milman's extension of Dvoretzky's theorem can be stated as

Theorem 1 *There is a function $c(\varepsilon) > 0$ such that for all $k \le c(\varepsilon)E(X)^2$, the space ℓ_2^k $(1+\varepsilon)$-embeds into X.*

By "U K-embeds into V" we mean here that there is an invertible linear transformation $A : U \to V' \subset V$ with $\|A\|\|A^{-1}\| \le K$. As a consequence (requiring additional arguments) one gets a closer relative of Dvoretzky's original theorem,

Theorem 2 *There is a function $c(\varepsilon) > 0$ such that for all $k \le c(\varepsilon)\log n$, ℓ_2^k $(1+\varepsilon)$-embeds into any normed space of dimension n.*

*Supported by the Israel Science Foundation

See [Dv] for the original theorem of Dvoretzky (in which the dependence of k on n is weaker), [Mi] for Milman's original work, [FLM] for expansions on Milman's method and [MS] and [Pi] for expository outlets of the subject (there are many others). The exposition in [Pi] is closer to our presentation here.

The dependence on n in Theorem 2 is known to be best possible (for ℓ_∞^n) but the dependence on ε is far from being understood. Gordon [Go] improved the dependence obtained from Milman's proof to $c(\varepsilon) \geq c\varepsilon^2$ for some universal $c > 0$. Another proof of that, following the (concentration) method of the proof from [Mi] is given in [Sc] and will be used here. As an upper bound for $c(\varepsilon)$ one gets $C/\log\frac{1}{\varepsilon}$ for some universal C. Indeed it is not hard to relate the smallest dimension n for which ℓ_2^k $1+\varepsilon$ - embeds into ℓ_∞^n to the size of a δ - net in S^{k-1}, for an appropriate δ. Using the known, and quite simple to attain, estimates on the size of such a net, we get that, for some universal $0 < c < C < \infty$, ℓ_2^k $1+\varepsilon$ - embeds into ℓ_∞^n if $k \leq \frac{c}{\log\frac{1}{\varepsilon}}\log n$ and, conversely, that if ℓ_2^k $1+\varepsilon$ - embeds into ℓ_∞^n then $k \leq \frac{C}{\log\frac{1}{\varepsilon}}\log n$.

In Section 2 of this note we improve the lower bound on $c(\varepsilon)$ by proving

Theorem 3 (First Main Theorem) *There is a constant $c > 0$ such that for all $n \in \mathbb{N}$ and all $\varepsilon > 0$, every n-dimensional normed space admits a subspace whose Banach–Mazur distance from ℓ_2^k is at most $1+\varepsilon$ and $k > \frac{c\varepsilon}{(\log\frac{1}{\varepsilon})^2}\log n$.*

In Section 3 we turn to the subject of embedding *subsets* of Euclidean spaces in normed spaces. A well known theorem of Johnson and Lindenstrauss [JL] asserts:

Theorem 4 *Let T be a k - point subset of an Euclidean space. Then, for every $\varepsilon > 0$, T $1+\varepsilon$ - Lipschitz embeds into ℓ_2^n with $n \leq \frac{C\log k}{\varepsilon^2}$.*

By "U K-Lipschitz embeds into V" (for U, V metric spaces) we mean here that there is an invertible map $f : U \to V' \subset V$ with the Lipschitz norm of f times the Lipschitz norm of f^{-1} at most K.

The proof of [JL] goes like that: Look at the set $S = \{\frac{t-s}{\|t-s\|}\ ;\ t, s \in T,\ t \neq s\}$ and find a *linear* map A from the Euclidean space containing T to ℓ_2^n, for the appropriate n, such that $1-\varepsilon \leq \|As\| \leq 1+\varepsilon$ for all $s \in S$. This will clearly do the job. In [JL] the map A is chosen randomly out of a class of orthogonal projections. Later it was also shown that one can use Gaussian or random ± 1 matrices for the same purpose. Recently, Klartag and Mendelson [KM] generalized this in two ways: First the linear map can be chosen out of a class of more general random matrices (the entries are identically distributed independent random variables with some prescribed tail behaviour) and secondly the estimate on n can be improved (sometimes). For a subset S of $\mathbb{R}^m$ put

$$E_S^* = \mathbb{E}\Big(\sup\{|\sum_{i=1}^n s_i g_i|;\ s = (s_1, \ldots, s_n) \in S\}\Big).$$

They proved that if S is a subset of S^{m-1} and $n \geq \frac{C(E_S^*)^2}{\varepsilon^2}$ (for some absolute C) then there is a linear $A : \mathbb{R}^m \to \ell_2^n$ with $1-\varepsilon \leq \|As\| \leq 1+\varepsilon$ for all $s \in S$. This easily implies the Johnson–Lindenstrauss result.

When we saw this result we noticed that, *for Gaussian matrices*, it follows easily from the method of [Sc] plus the statement of Talagrand's majorizing measure theorem ([Ta1] for the original theorem, [LT] for an expository outlet). Moreover, the method of [Sc] allows to give a good embedding theorem for S as above *in a general normed space.*

Theorem 5 (Second Main Theorem) *Let X be normed space and let T be a subset of S^{m-1}. Then, for every $\varepsilon > 0$, if $E_T^* \leq c\varepsilon E(X)$, there is a linear operator $A : \mathbb{R}^m \to X$ with*

$$1 - \varepsilon \leq \|At\| \leq 1 + \varepsilon$$

for all $t \in T$. $c > 0$ is a universal constant.

Since $E(\ell_2^n) \sim \sqrt{n}$, this generalizes the Gaussian case of [KM]. Note also that, taking $T = S^{m-1}$ and X general, we get back Theorem 1 with the best known dependence on ε.

2 The dependence on ε in Dvoretzky's Theorem

Lemma 1 *Let $\|\cdot\|$ be a norm on $\mathbb{R}^N$ satisfying $\|\cdot\| \leq |\cdot|$ and let $e_1, \ldots, e_n$ be an orthonormal sequence in $\mathbb{R}^N$ satisfying $\|x_i\| \geq 1/2$ for all i and*

$$\mathbb{E}\Big(\|\sum_{i=1}^{n} g_i e_i\| \Big) \leq L\sqrt{\log n}. \tag{2}$$

Then, for all disjoint $\sigma_1, \ldots, \sigma_{\lfloor\sqrt{n}\rfloor} \subset \{1, \ldots, n\}$ with $|\sigma_j| = \lfloor\sqrt{n}\rfloor$ for all j, there is a subset $J \subset \{1, \ldots, \lfloor\sqrt{n}\rfloor\}$ of cardinality at least $\frac{\sqrt{n}}{2}$ and there are $\{x_j\}_{j\in J}$ with x_j supported on σ_j such that $\|x_j\| = 1$ for all $j \in J$ and

$$\mathbb{E}\Big(\|\sum_{j\in J} r_j x_j\| \Big) \leq 80L.$$

Proof: A well known convexity argument for the first inequality and a standard estimate for the second imply that for all j

$$\mathbb{E}\Big(\|\sum_{i\in\sigma_j} g_i e_i\| \Big) \geq \frac{1}{2}\mathbb{E}(\max_{i\in\sigma_j} |g_i|) \geq \frac{1}{20}\sqrt{\log n}.$$

Since $\{x_i\}_{i\in\sigma_j} \to \|\sum_{i\in\sigma_j} x_i e_i\|$ is 1-Lipschitz (with respect to the ℓ_2 norm) we get from the standard concentration inequalities for Gaussian measures (see e.g. page 140 in [MS]) that

$$P\Big(\|\sum_{i\in\sigma_j} g_i e_i\| \leq \frac{1}{40}\sqrt{\log n} \Big) \leq e^{-\frac{\log n}{80}}.$$

It follows that, for $n \geq 2^{80}$, $P\Big(\|\sum_{i\in\sigma_j} g_i e_i\| \leq \frac{1}{40}\sqrt{\log n}\Big) \leq \frac{1}{2}$ for all j and, since these events are independent when j ranges over $1,\ldots,\lfloor\sqrt{n}\rfloor$, with probability at least $\frac{1}{2}$ there is a subset $J \subset \{1,\ldots,\lfloor\sqrt{n}\rfloor\}$ with $|J| \geq \frac{\lfloor\sqrt{n}\rfloor}{2}$ such that $\|\sum_{i\in\sigma_j} g_i e_i\| > \frac{1}{40}\sqrt{\log n}$ for all $j \in J$. Denote the event that such a J exists by A. Let $\{r_j\}_{j=1}^{\lfloor\sqrt{n}\rfloor}$ be a Rademacher sequence independent of the original Gaussian sequence. We get that

$$
\begin{aligned}
L\sqrt{\log n} \geq\ & \mathbb{E}_g\left(\|\textstyle\sum_{j=1}^{\lfloor\sqrt{n}\rfloor}\sum_{i\in\sigma_j} g_i e_i\|\right) = \mathbb{E}_r\mathbb{E}_g\left(\|\textstyle\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} r_j \sum_{i\in\sigma_j} g_i e_i\|\right) \\
\geq\ & \mathbb{E}_r\mathbb{E}_g\left(\|\textstyle\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} r_j \sum_{i\in\sigma_j} g_i e_i\|\mathbf{1}_A\right) = \frac{1}{2}\mathbb{E}_g\left(\left(\mathbb{E}_r\|\textstyle\sum_{j=1}^{\lfloor\sqrt{n}\rfloor} r_j \sum_{i\in\sigma_j} g_i e_i\|\right)\Big/A\right).
\end{aligned}
$$

It follows that for some $\omega \in A$, there exists a $J \subset \{1,\ldots,\lfloor\sqrt{n}\rfloor\}$ with $|J| \geq \frac{\lfloor\sqrt{n}\rfloor}{2}$ such that putting $\bar{x}_j = \sum_{i\in\sigma_j} g_i(\omega)e_i$, $\|\bar{x}_j\| \geq \frac{1}{40}\sqrt{\log n}$ and

$$
\mathbb{E}_r\left(\|\sum_{j\in J} r_j\bar{x}_j\|\right) \leq 2L\sqrt{\log n}.
$$

Take $x_j = \bar{x}_j/\|\bar{x}_j\|$. ■

Corollary 1 *With the assumptions of Lemma 1 there is a subspace of* $(\text{span}\{e_i\}_{i=1}^n, \|\cdot\|)$ *of dimension* $k \geq \frac{n^{1/4}}{CL}$ *which is* CL*-isomorphic to* ℓ_∞^k*.* C *is a universal constant.*

This follows immediately from Lemma 1 and Theorem 4.1 of [AM]. See also [Ta2] for a simpler proof of the result from [AM].

Remark 1 By starting with sets σ_j of size n^δ instead of $\sqrt{n}$, one easily gets a similar conclusion to that of Lemma 1 with $|J| \geq n^{1-\delta}$ and a constant C_δ depending on δ instead of 80. Consequently we get a strengthening of Corollary 1

Corollary 2 *With the assumptions of Lemma 1, for each* $\delta > 0$ *there is a constant* C_δ*, depending only on* δ*, and there is a subspace of* $(\text{span}\{e_i\}_{i=1}^n, \|\cdot\|)$ *of dimension* $k \geq \frac{n^{\frac{1}{2}-\delta}}{C_\delta L}$ *which is* $C_\delta L$*-isomorphic to* ℓ_∞^k*.*

Corollary 3 *With the assumptions of Lemma 1, for any* $0 < \varepsilon < 1$ *there is a subspace of* $(\text{span}\{e_i\}_{i=1}^n, \|\cdot\|)$ *of dimension* $k \geq cn^{\frac{c\varepsilon}{\log L}}$ *which is* $1+\varepsilon$*-isomorphic to* ℓ_∞^k*.* $c > 0$ *is a universal constant.*

This follows from Corollary 1 and a result of James. The argument is also reproduced in [AM].

Theorem 6 *There is a constant $c>0$ such that for all $n\in\mathbb{N}$ and all $0<\varepsilon<1$, every n-dimensional normed space admits a subspace whose Banach-Mazur distance from ℓ_2^k is at most $1+\varepsilon$ and $k>\frac{c\varepsilon}{(\log\frac{1}{\varepsilon})^2}\log n$.*
Equivalently, every symmetric convex body in $\mathbb{R}^n$ admits a k-dimensional section containing an Euclidean ball and contained in $1+\varepsilon$ times that ball where $k>\frac{c\varepsilon}{(\log\frac{1}{\varepsilon})^2}\sqrt{\log n}$.

Proof: We start with the setup of the proof of Theorem 1 as can be found for example in [MS]. Since the first statement in Theorem 6 is invariant under linear transformation we may assume that the normed space in question is $X=(\mathbb{R}^n,\|\cdot\|)$ where S^{n-1} is the ellipsoid of maximal volume contained in the unit ball of X. It follows from the Dvoretzky–Rogers Lemma that there is an orthonormal basis $e_1,\dots,e_n$ of $\mathbb{R}^n$ with $\|e_i\|\ge\frac{1}{2}$ for $i=1,\dots,\lfloor\frac{n}{2}\rfloor$. Note also that $\|\cdot\|\le|\cdot|$. Denote $E=\mathbb{E}\Big(\|\sum_{i=1}^n g_ie_i\|\Big)$ then Theorem 1 states that X admits a subspace whose Banach–Mazur distance from ℓ_2^k is at most $1+\varepsilon$ and $k>c\varepsilon^2E^2$ (more precisely, Milman's argument as presented in [FLM] only gives $k>c\frac{\varepsilon^2}{\log\frac{1}{\varepsilon}}E^2$. Gordon [Go] improved the dependence on ε to ε^2; see also [Sc] for another proof - more on that proof in the next section).

If $\varepsilon^2E^2\ge\frac{\varepsilon}{(\log\frac{1}{\varepsilon})^2}\log n$ we are thus done, so we may assume that

$$\mathbb{E}\Big(\|\sum_{i=1}^{\lfloor\frac{n}{2}\rfloor}g_ie_i\|\Big)\le E\le\frac{1}{\varepsilon^{1/2}\log\frac{1}{\varepsilon}}\sqrt{\log n}.$$

Apply now Corollary 3 to get a subspace of X of dimension $m\ge cn^{\overline{\frac{c\varepsilon}{\log(\varepsilon^{-1/2}(\log\frac{1}{\varepsilon})^{-1})}}}$ which is $1+\varepsilon$ - isomorphic to ℓ_∞^m. ℓ_∞^m contains a subspace of dimension at least $k=\frac{c}{\log\frac{1}{\varepsilon}}\log m$ which is $1+\varepsilon$ - isomorphic to an Euclidean space, for some universal constant $c>0$. (This is well known, here is the outline of the argument: Let $\{x_i\}_{i=1}^m$ be an ε-net on S^{k-1} of cardinality $m=\lfloor(\frac{3}{\varepsilon})^k\rfloor$ and consider the embedding $T:\ell_2^k\to\ell_\infty^m$ given by $Tx=(\langle x,x_i\rangle)_{i=1}^m$.) It follows that X contains a subspace of dimension k at least $\frac{c}{\log\frac{1}{\varepsilon}}\frac{c\varepsilon}{\log(\varepsilon^{-1/2}(\log\frac{1}{\varepsilon})^{-1})}\log n\ge\frac{c'\varepsilon}{(\log\frac{1}{\varepsilon})^2}\log n$ which is $(1+\varepsilon)^2$ - isomorphic to ℓ_2^k. This concludes the proof of the first assertion (since $(1+\varepsilon)^2\le1+3\varepsilon$ for $0<\varepsilon<1$). That the second, geometric, assertion of the theorem follows from the first is well known and easily follows from the fact that any $2m$ - dimensional ellipsoid in $\mathbb{R}^{2m}$ admits an m - dimensional central section which is an Euclidean ball. ■

3 Embedding subsets of Euclidean space in normed spaces

Here we bring a joint generalization of the Johnson–Lindenstrauss Lemma (Theorem 4) concerning Lipschitz embedding of subsets of Euclidean space in a low dimensional Euclidean spaces and of Milman's version of Dvoretzky's Theorem (Theorem 1) concerning embedding Euclidean spaces in general normed spaces.

Recall that given a normed space X we denote

$$E(X) = \sup\{\mathbb{E}\Big(\|\sum_{i=1}^{n} g_i u(e_i)\|_X\Big)\ ;\ n \in \mathbb{N},\ u : \ell_2^n \to X, \|u\| = 1\}$$

and that given a bounded subset T of $\mathbb{R}^m$ we denote

$$E_T^* = \mathbb{E}\Big(\sup\{|\sum_{i=1}^{m} t_i g_i|;\ t = (t_1, \dots, t_m) \in T\}\Big).$$

Note that letting $\|x\| = \sup\{|\langle x, t\rangle|;\ t \in T\}$, and $X = (\mathbb{R}^m, \|\cdot\|)$, $E_T^* \le E(X)$.

Theorem 7 *Let X be a finite dimensional normed space and let T be a subset of S^{m-1}. Then, for every $\varepsilon > 0$, if $E_T^* \le c\varepsilon E(X)$, there is a linear operator $A : \mathbb{R}^m \to X$ with*

$$1 - \varepsilon \le \|At\| \le 1 + \varepsilon$$

for all $t \in T$. $c > 0$ is a universal constant.

Note that this is a joint generalization of Milman's version of Dvoretzky's Theorem (with the best dependence on ε) and a generalization of the Johnson-Lindenstrauss lemma: If $T = S^{n-1}$ we get the first. If T is general and $X = \ell_2^k$ with $k \ge C\varepsilon^{-2}(E_T^*)^2$ we get the recent generalization of Klartag and Mendelson to the Johnson-Lindenstrauss lemma (in the Gaussian case).

One can get a conclusion similar to that of Theorem 7 by using first the special case $X = \ell_2^k$ and then Theorem 1 for embedding ℓ_2^k in X but then the dependence of ε will be worth.

One gets for example from Theorem 7 that any n-points set in a Hilbert space Lipschitz embeds in ℓ_1^k for k of order $\frac{\log n}{\varepsilon^2}$. Was that known previously?

Proof of Theorem 7: The proof follows that of the main theorem of [Sc] with a twist at the end. We may assume that X is finite dimensional, say $X = (\mathbb{R}^n, \|\cdot\|)$, that the sup in the definition of $E(X)$ is attained for the same n and (by applying an isometry) for u being the identity map. Put $E = E(X) = \mathbb{E}\Big(\|\sum_{i=1}^n g_i e_i\|\Big)$. Let $\{g_{ij}\}_{i=1}^{m}{}_{j=1}^{n}$ be independent standard Gaussian variables on some probability space and for each ω in this probability space and $a = (a_1, \dots, a_m)$ in $\mathbb{R}^m$ define

$$B_\omega(a) = \sum_{i=1}^{m} a_i \sum_{j=1}^{n} g_{i,j} e_j. \tag{3}$$

We may assume that T is not empty. Let $t_0 \in T$ and for $a \in S^{m-1}$ put

$$H_\omega(a) = \|B_\omega(a)\| - \|B_\omega(t_0)\|.$$

Note that, for all $a \in S^{m-1}$, $\mathbb{E}H_\omega(a) = 0$. The next lemma was proved in [Sc]; we shall repeat the proof (and slightly extend it) bellow since [Sc] may be hard to find.

Lemma 2 *For some absolute constant C the process $\{H_\omega(a)\}_{a\in S^{m-1}}$ is subgaussian with respect to the metric $d(a,b)=C\|a-b\|_2$. i.e., for all $s>0$,*

$$P(|H_\omega(a)-H_\omega(b)|>s)\le 2\exp\Big(\frac{-s^2}{C\|a-b\|_2^2}\Big).$$

Consider also the Gaussian process

$$G(a)=G_\omega(a)=\sum_{i=1}^m a_i g_i$$

whose corresponding metric is $(\mathbb{E}(G(a)-G(b))^2)^{1/2}=\|a-b\|_2$. By the majorizing measure theorem in its comparison form (see e.g., Theorem 12.16 in [LT]), for some absolute constant K,

$$\begin{aligned}\mathbb{E}(\sup_{t\in T}|\|B_\omega(t)\|-\|B_\omega(t_0)\||) &\le \mathbb{E}(\sup_{t\in T}H_\omega(t))+\mathbb{E}(\sup_{t\in T}-H_\omega(t))\\ &\le K\mathbb{E}(\sup_{t\in T}G(t))+K\mathbb{E}(\sup_{t\in T}-G(t))\le 2KE_T^*.\end{aligned}$$

It follows that, if $8KE_T^*\le\varepsilon E$, then $\mathbb{E}(\sup_{t\in T}|\|B_\omega(t)\|-\|B_\omega(t_0)\||)\le\varepsilon E/4$ and thus, with probability at least 1/2, there is an ω for which

$$|\|B_\omega(t)\|-\|B_\omega(t_0)\||\le\varepsilon E/2\qquad\text{for all } t\in T. \tag{4}$$

Also, since the function $(a_1,\dots,a_n)\to\|\sum_{j=1}^n a_je_j\|$ is 1-Lipschitz,

$$P\Big(|\|B_\omega(t_0)\|-E|>\frac{\varepsilon}{2}E\Big)=P\Big(|\|\sum_{j=1}^n g_je_j\|-E|>\frac{\varepsilon}{2}E\Big)\le e^{-c'\varepsilon^2\mathbb{E}^2}, \tag{5}$$

for some absolute $c'>0$. Since E_T^* is at least 1, we may assume that εE is large enough so that the right hand side of (5) is smaller than 1/2. It follows that, with probability larger than 1/2,

$$(1-\frac{\varepsilon}{2})E\le\|B_\omega(t_0)\|\le(1+\frac{\varepsilon}{2})E.$$

This together with (4) shows that there is an ω for which

$$(1-\varepsilon)E\le\|B_\omega(t)\|\le(1+\varepsilon)E\qquad\text{for all } t\in T.$$

Take $A=B_\omega/E$. ■

We now state and prove a slightly extended version of Lemma 2. With the definition of $B(a)=B_\omega(a)$ as in (3), extend the definition of $H(a)=H_\omega(a)$ to all $a\in\mathbb{R}^m$ by

$$H(a)=H_\omega(a)=\|B_\omega(a)\|-\|a\|_2\|B_\omega(t_0)\|.$$

Note that $H(a)$ has mean zero for each $a\in\mathbb{R}^m$.

Lemma 3 *For some universal constant* C*,*

$$P(|H_\omega(a) - H_\omega(b)| > s) \le 6\exp\Big(\frac{-s^2}{C\|a-b\|_2^2}\Big)$$

for all $a, b \in \mathbb{R}^m$ *and all* $s > 0$*.*

Proof: First assume that $\|a\|_2 = \|b\|_2$. Note that this case is all that is needed for the proof of Theorem 7. Put $c = \frac{a+b}{2}$ and notice that, since $b-a$ and c are orthogonal, $B(\frac{b-a}{2})$ is independent of $B(c)$. Fix an $x \in \mathbb{R}^m$ and consider the function $f : \mathbb{R}^{mn} \to \mathbb{R}$ given by

$$f_{a-b}(\{\alpha_{ij}\}) = \Big\| x + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}(a_i - b_i)\alpha_{ij}e_j \Big\|.$$

This is a Lipschitz function with constant $\|a-b\|_2/2$. Denote its expectation with respect to the canonical gaussian measure on $\mathbb{R}^{mn}$ by E_x, then by the concentration inequality for Gaussian measures (see e.g., page 140 in [MS]),

$$P\Big(\Big|\Big\| x + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}(a_i - b_i)g_{ij}e_j \Big\| - E_x\Big| > s\Big) \le 2\exp(-cs^2/\|a-b\|_2^2)$$

for all $s > 0$ and some absolute $c > 0$ The same is true for the function f_{b-a} (with the *same* E_x). It follows that, conditioning on $B(\frac{a+b}{2}) = x$,

$$\begin{aligned} P(|H(a) - H(b)| > s) &= \\ &= P\Big(\Big|\Big\| x + \tfrac{1}{2}\textstyle\sum_{i=1}^{m}\sum_{j=1}^{n}(a_i - b_i)g_{ij}e_j \Big\| - \Big\| x + \tfrac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{n}(b_i - a_i)g_{ij}e_j \Big\|\Big| > s\Big) \\ &\le 2P\Big(\Big|\Big\| x + \tfrac{1}{2}\textstyle\sum_{i=1}^{m}\sum_{j=1}^{n}(a_i - b_i)g_{ij}e_j \Big\| - E_x\Big| > s/2\Big) \\ &\le 4\exp(-cs^2/\|a-b\|_2^2) \end{aligned}$$

and thus the same inequality holds also without the conditioning.

Now consider the general case and assume for example that $\|a\|_2 < \|b\|_2$. Denote $\bar{b} = \frac{\|a\|_2}{\|b\|_2}b$. Then,

$$H(b) - H(\bar{b}) = \Big(1 - \frac{\|a\|_2}{\|b\|_2}\Big)\|B(b)\| - (\|b\|_2 - \|a\|_2)E = (\|b\|_2 - \|a\|_2)H(b/\|b\|_2).$$

It follows that

$$\begin{aligned} P(|H(b) - H(\bar{b})| > s) = &\quad P(|H(b/\|b\|_2)| > s/(\|b\|_2 - \|a\|_2)) \\ \le &\quad 2\exp(-cs^2/(\|b\|_2 - \|a\|_2))^2 \le 2\exp(-cs^2/\|b-\bar{b}\|_2^2), \end{aligned}$$

and thus,

$$P(|H(a)-H(b)|>s) \leq P(|H(a)-H(\bar{b})|>s/2)+P(|H(b)-H(\bar{b})|>s/2)$$
$$\leq 4\exp(-cs^2/\|a-\bar{b}\|_2^2)+2\exp(-cs^2/\|b-\bar{b}\|_2^2).$$

Since $\|a-b\|_2 \geq \max\{\|a-\bar{b}\|_2, \|b-\bar{b}\|_2\}$ we get the desired conclusion. ■

References

[AM] Alon, N.; Milman, V. D., Embedding of l_∞^k in finite-dimensional Banach spaces. Israel J. Math. 45 (1983), no. 4, 265–280.

[Dv] Dvoretzky, A., Some results on convex bodies and Banach spaces. 1961 Proc. Internat. Sympos. Linear Spaces (Jerusalem, 1960) pp. 123–160 Jerusalem Academic Press, Jerusalem; Pergamon, Oxford.

[FLM] Figiel, T.; Lindenstrauss, J.; Milman, V. D., The dimension of almost spherical sections of convex bodies. Acta Math. 139 (1977), no. 1-2, 53–94.

[Go] Gordon, Y., Some inequalities for Gaussian processes and applications. Israel J. Math. 50 (1985), no. 4, 265–289.

[JL] Johnson, W. B.; Lindenstrauss, J., Extensions of Lipschitz mappings into a Hilbert space. Conference in modern analysis and probability (New Haven, Conn., 1982), 189–206, Contemp. Math., 26, Amer. Math. Soc., Providence, RI, 1984.

[KM] Klartag B.; Mendelson, S., Empirical processes and random projections.

[LT] Ledoux, M.; Talagrand, M., Probability in Banach spaces. Isoperimetry and processes. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 23. Springer-Verlag, Berlin, 1991.

[Mi] Milman, V. D., A new proof of A. Dvoretzky's theorem on cross-sections of convex bodies. (Russian) Funkcional. Anal. i Priložen. 5 (1971), no. 4, 28–37.

[MS] Milman, V. D. and Schechtman, G., Asymptotic theory of finite-dimensional normed spaces, Lecture Notes in Mathematics, 1200, Springer-Verlag, Berlin, 1986.

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

[Sc] Schechtman, G., A remark concerning the dependence on ϵ in Dvoretzky's theorem. Geometric aspects of functional analysis (1987–88), 274–277, Lecture Notes in Math., 1376, Springer, Berlin, 1989.

[Ru] Rudelson, M., Estimates of the weak distance between finite-dimensional Banach spaces. Israel J. Math. 89 (1995), no. 1-3, 189–204.

[Ta1] Talagrand, M., Regularity of Gaussian processes. Acta Math. 159 (1987), no. 1-2, 99–149.

[Ta2] Talagrand, M., Embedding of l_k^∞ and a theorem of Alon and Milman. Geometric aspects of functional analysis (Israel, 1992–1994), 289–293, Oper. Theory Adv. Appl., 77, Birkhauser, Basel, 1995.

Gideon Schechtman
Department of Mathematics
Weizmann Institute of Science
Rehovot, Israel
E-mail: gideon.schechtman@weizmann.ac.il