Алгебра и анализ
Том. 19 (2007), №1

St. Petersburg Math. J.
S 1061-0022(XX)0000-0

USING RADEMACHER PERMUTATIONS TO REDUCE RANDOMNESS

S. ARTSTEIN-AVIDAN AND V. D. MILMAN

Dedicated to Professor V. A. Zalgaller on the occasion of his 85th birthday

Abstract. It is shown how a special family of unitary operators, called the Rademacher permutations and related to the Clifford algebra, can be used to reduce the level of randomness in several results in asymptotic geometric analysis.

§1. Introduction: Derandomization

In asymptotic geometric analysis, we investigate structures and phenomena that occur in normed spaces of high dimensions. Enormously rich structures have already been discovered (see, e.g., the survey [GiM]), and there is a lot of ongoing research in this direction. However, the existence of essentially all these remarkable objects and structures was proved by probabilistic arguments. Such an object can be, for example, special subspaces of a normed space that are close to Euclidean, or special vectors that create a zigzag body (to be defined below) and about which we know that, when chosen independently at random, they will possess certain joint properties. At the same time, we know next to nothing about how to construct such objects explicitly, or how to demonstrate such structures explicitly. Questions in this spirit were asked already in the beginning of the 80's (see, e.g., [M]) and as time passes, the lack of information increasingly bothers experts.

Besides the natural curiosity about the reasons behind the lack of explicit constructions, and about why we cannot locate even one such object when we know how to prove their existence with probability close to one exponentially (in dimension), there is another, more pragmatic reason for the study of explicit constructions. We know, for example, that the explicit constructions of expanders (see, e.g., [AS]), which were discovered years after their existence was established by probabilistic arguments, led to a very beautiful theory and many important applications. In a more geometric context, the first explicit construction of a projection of a section of the simplex in $\mathbb{R}^n$ of dimension proportional to n and ε-close to the Euclidean ball, was discovered very recently by Ben-Tal and Nemirovski, see [BN], which gave an unexpectedly good dependence on the parameters involved, namely, the dimension n and the level of approximation ε. Thus, it may be expected that finding explicit constructions will even provide better existence results, in some cases.

The definition of an 'explicit construction' involves the computing time needed to generate the corresponding structure, and may depend on the computation model chosen. We leave this definition vague, since in all the specific cases to be treated below it will be clear what is and what is not explicit. For example, an exhaustion algorithm, which consists of trying out all possible structures (say, in a discrete setting) until bumping into one that possesses the properties sought (which will be found, eventually, since probabilistic arguments show that such a structure exists), is not viewed as an explicit

©2004 American Mathematical Society

construction because, typically, there will be exponentially many structures to examine. Of course, when one is allowed to use randomness, this searching procedure can usually be made very efficient, but in our setting avoiding randomness is the main goal.

In this paper we continue the line of investigation we began in [ArM]. To understand the structures behind explicitness, first we shall go "half-way". We do not eliminate randomness completely, but we study how much one may reduce it. On our way to such a reduction, we found some new structures, which are undoubtedly of independent interest. One such structure, and what may be achieved by using it, is presented in this paper.

The amount of randomness needed for a certain result is typically measured in the number of random *bits* (or coin-flips) needed. Another parameter that plays a role is the final probability with which a 'good structure' is arrived at. Since in asymptotic geometric analysis we often deal with continuous measures, we would like to measure randomness by the amount of random *Gaussian* "coin flips", or by the amount of random sphere vectors (which is n — the dimension — times the number of Gaussian "coin flips", since a vector of independent Gaussian entries is, after normalization, uniformly distributed on the sphere). However, when this is done, several issues occur. First, it can be seen that a random Gaussian can be decomposed into two independent random Gaussians, and therefore such a "measure" of the amount of randomness does not, formally, make sense. Second, one cannot usually expect to work with the exact uniform Haar measure on a compact but infinite group (such as S^{n-1} or $O(n)$), or with the exact Gaussian measure, but rather with some (usually finite) approximation of it.

To solve this it is natural to discuss also "distorted" uniform measures, which forces us to consider approximate-uniform measures, and we may use those to measure randomness. There are many different ways to define approximate-uniform measures, and we chose a way that is most useful for our goal below, but not necessarily for other goals. In Subsection 3.3, we describe this notion and prove a generalization to the approximate-uniform setting of one theorem from [ArM]. However, we remark that, usually, the intuition coming from counting truly-random Gaussians which are used "routinely", works well. Thus, in all other sections, we shall work with truly-uniform objects and not with the approximate-uniform, but the same generalization applies. Therefore, the results can be interpreted as reduction of randomness results, besides their standard interpretation as structural results in asymptotic geometric analysis. A relationship between the number of approximate-Gaussian coin flips and the number of random bits also exists, because an approximate-Gaussian can be generated by using random signs, but we do not pursue this direction in this paper.

In [ArM], we described several results about reduction of randomness for geometric constructions in high dimensions. In the main part of this paper we describe a new tool, which helps further reduce randomness, where the generic random object is either random uniform sphere vectors or random unitary matrices distributed uniformly in $O(n)$ (but generalization to the approximate-uniform setting is possible). The method works best in $\mathbb{R}^n$ when n is a power of 2, and adjustments can usually be made, depending on the specific issue in question, to get results in general dimension. Our main tool, which we call Rademacher permutations, is related to the Clifford Algebra. Usually, it improves randomness reduction results by a logarithmic factor.

Notation. Let S^{n-1} denote the Euclidean sphere of radius 1 in $\mathbb{R}^n$ equipped with the normalized Haar measure σ, and let D_n denote the Euclidean ball. The Euclidean norm of a vector x will be denoted by $|x|$, so that $D_n = \{x \in \mathbb{R}^n : |x| \leq 1\}$. For a symmetric convex body $K \subset \mathbb{R}^n$, we denote by $\|\cdot\|_K$ the norm whose unit ball is K. We put $M = M(K) = \int_{S^{n-1}} \|u\|_K d\sigma(u)$, and denote $M^*(K) = M(K^\circ)$ where K° is the dual

body for K, that is, $K^\circ = \{u : \sup_{x\in K}\langle u, x\rangle \le 1\}$. The quantity $M^*(K)$ is called half the mean width of K, because $M^*(K) = \int_{S^{n-1}} \sup_{x\in K}\langle u, x\rangle d\sigma(u)$. Usually, we employ d for the half-diameter (or radius) of K, and b for the half-diameter of its dual; these are the smallest possible constants for which $\frac{1}{b}D_n \subseteq K \subseteq dD_n$. We let $O(n)$ stand for the group of orthogonal $(n \times n)$-matrices endowed with the normalized Haar measure. The distance on $O(n)$ is determined by the Hilbert–Schmidt norm. Sometimes, we ignore integer values to avoid complicating the notation. We use the letters c, c_i, c', C, C_i, C' to denote universal constants, which may be different in different parts of the paper. When a constant depends on some other parameter, we indicate this parameter in parentheses (e.g., $C_0(\varepsilon)$ is a constant depending only on ε).

§2. Rademacher permutations

Throughout this section, we assume that $n = 2^k$ for some integer k. We shall construct $k = \log n$ orthogonal transformations, each being a signed permutation, with the property that for any nonzero vector x the vectors

$$U_0 x, U_1 x, \dots, U_{(k-1)} x$$

are mutually orthogonal.

In fact, the operators will be constructed so that $U_0 = \mathrm{Id}$, $U_j^t = -U_j$ (and so $U_j^2 = -I$) for $j \ge 1$, and $U_i U_j = -U_j U_i$ for $i, j \ge 1$. This already implies that the vectors $U_j x$ are mutually orthogonal.

A way to define these $\log n$ orthogonal signed-permutation operators is as follows. For a vector x with n entries, we index the entries in the binary form. Since $n = 2^k$, each index is given by a word of length k, that is, a word with k digits, each being 0 or 1. The jth permutation will be performed by switching each two entries whose index differs at the jth digit only. Next, we need to define the signs. For this we use the Rademacher sequence, which is defined as follows: r_j is a vector with sign at a given index depending only on the jth digit of the index. The sign is $+$ if the jth digit of the index is 1, and $-$ if the jth digit is 0. (So, if the binary digits are replaced by $-$ and $+$ signs, and the indices are written in the left to right order, as the columns of a matrix, then the Rademachers are simply the rows of this matrix.) We introduce the operator U_j by taking the jth permutation defined above and multiplying it by $r_1 \cdots r_j$ (coordinatewise multiplication).

Thus, the Rademachers are defined as

$$\begin{aligned}
r_1 &= (-,+,-,+,-,+,-,+,-\cdots),\\
r_2 &= (-,-,+,+,-,-,+,+,-\cdots),\\
r_3 &= (-,-,-,-,+,+,+,+,-\cdots),\\
r_4 &= (-,-,-,-,-,-,-,-,+\cdots),\\
&\vdots
\end{aligned}$$

and our sequences of signs as

$$\begin{aligned}
r_1 r_2 &= (+,-,-,+,+,-,-,+,+\cdots),\\
r_1 r_2 r_3 &= (-,+,+,-,+,-,-,+,-\cdots),\\
r_1 r_2 r_3 r_4 &= (+,-,-,+,-,+,+,-,-\cdots),\\
&\vdots
\end{aligned}$$

The signed permutations we finally use are

$$\begin{aligned}
U_0(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,\ldots) &= (\ x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8,\ldots),\\
U_1(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,\ldots) &= (-x_2,\ x_1,-x_4,\ x_3,-x_6,\ x_5,-x_8,\ x_7,\ldots),\\
U_1(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,\ldots) &= (\ x_3,-x_4,-x_1,\ x_2,\ x_7,-x_8,-x_5,\ x_6,\ldots),\\
U_2(x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,\ldots) &= (-x_5,\ x_6,\ x_7,-x_8,\ x_1,-x_2,-x_3,\ x_4,\ldots),\\
&\vdots
\end{aligned}$$

Proof. We must show that $U_j = -U_j^t$ and $U_iU_j = -U_jU_i$ for $i \neq j$. The first property will follow if we prove that, for any two indices differing at the jth digit, the entries of $r_1 \cdots r_j$ are opposite. But this is true, because the entries of $r_1 \cdots r_{j-1}$ will be the same for both indices, and the entry of r_j will be different. For the second property, first observe that both U_iU_j and U_jU_i are signed permutations. Moreover, these permutations are simply exchanging of couples: switching two entries if their indices are different at the ith and the jth digit exactly. This already shows that the entries of U_iU_j and U_jU_i are equal up to sign. To compare the signs, we consider a specific entry $(U_jU_ix)_k$. This is equal to $\pm(U_ix)_{k_1}$, where k and k_1 differ at the jth digit only, and the sign is the sign in the kth entry of $r_1 \cdots r_j$. In its turn, this is equal to $\pm(x)_{k'}$, where k and k' differ at the jth and the ith digits only, and the sign is the product of the sign in the kth entry of $r_1 \cdots r_j$ and the sign in the k_1th entry of $r_1 \cdots r_i$.

On the other hand, the entry $(U_iU_jx)_k$ equals $\pm(x)_{k'}$, where the sign is the product of the sign in the kth entry of $r_1 \cdots r_i$ and the sign in the k_2th entry of $r_1 \cdots r_j$, where k_2 differs from k at the ith digit only.

Assume that $i < j$. All the indices we consider differ only at the ith and/or at the j^{th} digits, so that we may cancel out all the Rademachers that do not distinguish between them. We are left with comparing the product of the sign in the kth entry of r_ir_j and the sign in the k_1th entry of r_i, with the product of the sign in the kth entry of r_i and the sign in the k_2th entry of r_ir_j. Since k_1 and k_2 differ at the ith digit, and all of the rest cancels, we are left with opposite signs, which completes the proof. □

Remark. For $n = 2,\ 4,\ 8$ we can construct n such operators; in $\mathbb{R}^4$ the four signed permutations can be

$$\begin{aligned}
&(x_1,\ x_2,\ x_3,\ x_4),\\
&(x_2,-x_1,\ x_4,-x_3),\\
&(x_3,-x_4,-x_1,\ x_2),\\
&(x_4,\ x_3,-x_2,-x_1),
\end{aligned}$$

and in $\mathbb{R}^8$ we can take

$$\begin{aligned}
&(x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8),\\
&(x_2,-x_1,\ x_4,-x_3,-x_6,\ x_5,-x_8,\ x_7),\\
&(x_3,-x_4,-x_1,\ x_2,-x_7,\ x_8,\ x_5,-x_6),\\
&(x_4,\ x_3,-x_2,-x_1,-x_8,-x_7,\ x_6,\ x_5),\\
&(x_5,\ x_6,\ x_7,\ x_8,-x_1,-x_2,-x_3,-x_4),\\
&(x_6,-x_5,-x_8,\ x_7,\ x_2,-x_1,-x_4,\ x_3),\\
&(x_7,\ x_8,-x_5,-x_6,\ x_3,\ x_4,-x_1,-x_2),\\
&(x_8,-x_7,\ x_6,-x_5,\ x_4,-x_3,\ x_2,-x_1).
\end{aligned}$$

However, in general we cannot expect an essentially larger number of operators satisfying these properties than the amount $\log n$ we constructed above. The fact that one cannot

produce more than $2\log n + 2$ such operators in $O(n)$ is well known, and follows from the representation theory of the Clifford algebra (but can also be shown by using simple linear algebra only).

§3. Global Dvoretzky theorem

3.1. Background. Given a convex body $K \subset \mathbb{R}^n$, with half diameter d and half mean width M^*, the global Dvoretzky-type theorem, which first appeared in [BLM] in an implicit form, and later explicitly in [MS2], states that with high probability, the Minkowski average of $C(d/M^*)^2$ random rotations of the body K is isomorphic to a Euclidean ball of radius M^*. The constants of isomorphism depend on the choice of C, and so does the probability estimate. We state the theorem in the ε-isometric case, in its equivalent dual form.

Theorem 1. *For any $\varepsilon > 0$ there exist constants $c(\varepsilon)$ and $C(\varepsilon)$, depending only on ε, such that for any dimension n and any $K \subset \mathbb{R}^n$ satisfying $\frac{1}{b}D \subseteq K$ and $M = M(K)$, with probability $1 - e^{-c(\varepsilon)n}$, for the $N = C(\varepsilon)(\frac{b}{M})^2$ random uniformly distributed $V_1, \dots, V_N \in O(n)$ we have*

$$(1-\varepsilon)M|x| \le \frac{1}{N}\sum_{i=1}^{N} \|V_i x\|_K \le (1+\varepsilon)M|x| \tag{1}$$

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

The "level of randomness" in this result is $N = C(\frac{b}{M})^2$ random orthogonal transformations (a random orthogonal transformation can be constructed by using n^2 Gaussian "coin tosses", for example, by Gram–Schmidt orthogonalization of n Gaussian vectors).

This bound for the number of rotations involved is optimal up to constants, see [MS1]. Moreover, in [MS1] it was shown that this bound is optimal regardless of the way the rotations are constructed. Namely, also in the nonrandom setting, one cannot do with essentially smaller number of rotations, that is, one cannot do with less than $c(\frac{b}{M})^2$ rotations for some different constant c. However, the question about explicit construction of such operators (even for some simple norms) remains open.

In the paper [ArM] it was noted that the "level of randomness" in this result can be reduced to $N = C_3 + C_4\log(\frac{b}{M})^2$ random rotations, where C_3 and C_4 are constants depending only on $\varepsilon > 0$ (in [ArM] only the isomorphic version was stated, but the $(1+\varepsilon)$-isometric version admits precisely the same proof). The number of operators was of course larger, as it must be by the results of [MS1], but they were all constructed explicitly by using only N *random* operators. This was done with the help of iteration, and the resulting operators were products of randomly chosen ones. To formulate that result precisely, it is notationally convenient to denote by $V_i^0 = \mathrm{Id}$ the identity operator, and to put $V_i^1 = V_i$. In [ArM] we proved the following.

Theorem 2. *There exist universal constants c_0, c_1, C_2, C_3, and C_4 such that for every symmetric convex body $K \subset \mathbb{R}^n$ satisfying $\frac{1}{b}D \subseteq K$, and with $M = M(K)$, with probability greater than $1 - e^{-c_1 n}$, for the $N = C_3 + C_4\log(\frac{b}{M})^2$ random uniformly distributed $V_1, \dots, V_N \in O(n)$ we have*

$$c_0 M|x| \le \frac{1}{2^N}\sum_{\delta\in\{0,1\}^N} \Big\| \prod_{i=1}^{N} V_i^{\delta_i} x \Big\|_K \le C_2 M|x|$$

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

In fact, this is true with C_4 arbitrarily close to 1, but then, as $C_4 \to 1$, C_3 grows to infinity and $c_0 \to 0$.

Another way to state this result is as follows: *In the same notation as above, for* $\|x\|_1 = \|x\|_K$ *and* $\|x\|_j = \frac{1}{2}(\|x\|_{j-1} + \|V_j x\|_{j-1})$ *and for* N *defined in Theorem 2, with probability* $1 - e^{-c_1 n}$ *we have*

$$c_0 M|x| \leq \|x\|_N \leq C_2 M|x|.$$

Remark. As was remarked above, this result was not stated in [ArM], but it has a similar $(1+\varepsilon)$-isometric version. Also, notice that the number of final operators used here is not optimal, namely, we use $2^N = 2^{C_3}(\frac{b}{M})^{2C_4}$ operators.

The above result remains valid if we change the randomness notion from *uniform* on $O(n)$ to almost-uniform (see the definition in Subsection 3.3 below). While this might seem like a technical point, this is essential if we wish to estimate the "level of randomness" in a valid way. Indeed, as was explained in the Introduction, if one is allowed to use even only one truly uniform continuous random variable (or Gaussian), one may construct from it as many independent identically distributed copies as desired. This does not happen in the discrete, or approximate-uniform case. The definition of approximate-uniformity and the generalization of the above theorem to this case are given in Subsection 3.3.

3.2. Main theorems. In the main part of this paper, we improve the above results to

$$N = [C_5 \log(b/M)/\log\log n] + 1$$

random rotations for a universal C_5, and moreover, show that if $C_6(b/M)^2 < \log n$ for some universal C_6, then with the aid of only *one* random rotation we may build explicitly rotations that satisfy (1). The number of rotations we build will still be above the optimal order of $(b/M)^2$, but surprisingly close to it. For this, we use the Rademacher permutations defined in §2. In fact, we prove the following $(1+\varepsilon)$-isometric theorem.

Theorem 3. *For every* $0 < \varepsilon < 1$, *there exist constants* $1 < C_0(\varepsilon) < \infty$ *and* $0 < c_1(\varepsilon) < 1$ *depending only on* ε *and such that, for any dimension* $n = 2^{k_0}$ *with some integer* k_0 *and for any norm* $\|\cdot\|$ *defined on* $\mathbb{R}^n$ *with average* M *and maximum* b, *setting* $k = [C_0(b/M)^2]$ *we have the following: if* $k \leq \log n$, *then the probability of choosing a random rotation* V *such that the (explicit) rotations* $U_1, \ldots, U_k$ *defined in §2 satisfy*

$$(1-\varepsilon)M_0|x| \leq \frac{1}{k}\sum_{i=0}^{k-1} \|VU_i x\|_K \leq (1+\varepsilon)M_0|x|$$

for every $x \in \mathbb{R}^n$, *is greater than* $1 - e^{-c_1 n}$, *where* M_0 *is the Levy mean of* $\frac{1}{k}\sum_{i=1}^{k} \|Ve_i\|_K$ *as a function on* $O(n)$.

Here $c_1 = c_1(\varepsilon) = \ln(c/\varepsilon)$ and $C_0 = C_0(\varepsilon) = C\ln(1+1/\varepsilon)/\varepsilon^2$ for universal c, C. We remark that $M_0 < M < M_0 + c''b/\sqrt{n}$ for a universal c''; see, e.g., [MS1].

Theorem 4. *For any* $0 < \varepsilon < 1$, *there exist constants* $C_0(\varepsilon)$, $C_2(\varepsilon)$, *and* $c_3(\varepsilon)$ *such that for any dimension* $n = 2^k$ *with some integer* k *and any norm* $\|\cdot\|$ *defined on* $\mathbb{R}^n$, *we may use*

$$T = \frac{\log\left(C_0 \frac{(b/M)^2}{\log n}\right)}{\log\left(\frac{\log n}{C_2^2}\right)} + 1 = \frac{\log\left(\frac{C_0}{C_2^2}(b/M)^2\right)}{\log\left(\frac{\log n}{C_2^2}\right)}$$

random operators $V_1, V_2, \ldots, V_T$ distributed uniformly in $O(n)$ to build explicitly $N = C_2^{2T-2}(b/M)^2$ operators $W_1, \ldots, W_N$ for which, with probability greater than $1 - e^{-c_3 n}$,

$$(1-\varepsilon)M_0|x| \le \frac{1}{N}\sum_{k=1}^{N}\|W_k x\|_K \le (1+\varepsilon)M_0|x| \tag{2}$$

for every $x \in \mathbb{R}^n$. The operators W_k will be of the form $\prod_j V_j U_{i_j}$ with U_i as defined in §2.

Remark. Notice that the number of final rotations used is surprisingly close to the optimal one, namely, it is at most $N = C'(b/M)^{2+c'/\log\log n}$. The condition $n = 2^k$ is discussed in Subsection 3.7.

3.3. Approximate uniformity. This subsection contains a small detour from the main line of the paper, and is aimed at justifying the term "reduction of randomness", which we use throughout the text. As was remarked above and explained in the Introduction, measuring randomness is better done with some notion of approximate-uniform distributions rather than truly-uniform ones. In this subsection we would like to give one possible definition for approximate-uniform sphere vectors and for approximate-uniform rotations, and exemplify a generalization of a theorem to this case. The theorem we generalize is Theorem 2 from [ArM]. We use the following definition of approximate uniformity, which is not necessarily the most general, but is adequate for our needs.

Definition 5. 1. A random sphere vector $X : \Omega \to S^{n-1}$ is said to be an (ε, α)-approximate uniform sphere vector if for any set $A \subset S^{n-1}$ of measure $1/2$, and for any $t > \varepsilon$, the probability that X belongs to the t-neighborhood A_t of A is greater than $1 - e^{-\alpha t^2 n}$.

2. A random operator $X : \Omega \to SO(n)$ is an (ε, α)-approximate uniform rotation if for any set $A \subset O(n)$ of measure $1/2$, and for any $t > \varepsilon$, the probability that X belongs to the t-neighborhood (in the Hilbert–Schmidt norm) A_t of A is greater than $1 - e^{-\alpha t^2 n}$.

Remark. Notice that the second definition implies in particular that if X is an (ε, α)-approximate uniform rotation, then for any $y \in S^{n-1}$ the vector Xy is an (ε, α)-approximate random sphere vector. For some applications, such as the one below, this property suffices; however, for the approximate-uniform versions of the other theorems in this paper, the more restrictive definition above is needed.

Now we restate Theorem 2 in the case of approximate-uniform rotations (and in the ε-isometric case).

Theorem 6. *There exist universal constants c_0, C_3, and for any $0 < \alpha$ and $0 < \varepsilon < 1$ there exist constants $\varepsilon_0(\varepsilon, \alpha)$, $c_1(\varepsilon, \alpha)$, $C_2(\alpha)$, and $C_4(\varepsilon, \alpha)$, such that for any dimension $n > n_0(\alpha, \varepsilon) = (\varepsilon\alpha/c_0)^{c_0}$, for every symmetric convex body $K \subset \mathbb{R}^n$ satisfying $\frac{1}{b}D \subseteq K$, and with $M = M(K)$, for $N = C_4 + N_1 = C_4 + C_2 + C_3 \log(\frac{b}{M})$ random (ε_0, α)-approximate uniform rotations $V_1, \ldots, V_{N_1}$, $W_1, \ldots, W_{C_4} \in O(n)$ the following is true: with probability greater than $1 - e^{-c_1 n}$ we have*

$$(1-\varepsilon)M|x| \le \frac{1}{2^N}\sum_{j=1}^{C_4}\sum_{\delta\in\{0,1\}^{N_1}}\Big\|\prod_{i=1}^{N_1} V_i^{\delta_i} W_j x\Big\|_K \le (1+\varepsilon)M|x|$$

for every $x \in \mathbb{R}^n$. One may take $\varepsilon_0(\varepsilon, \alpha) = \varepsilon\sqrt{\alpha}/C_0$, $c_1(\varepsilon, \alpha) = \min(\alpha, \ln(11/\varepsilon))/C_0$, $C_2(\alpha) = C_0(1 + \ln(\alpha))$, and $C_4(\varepsilon, \alpha) = C_0 \frac{1}{\varepsilon^2\alpha^2}\ln(\frac{11}{\varepsilon})$, where C_0 is universal.

Remark. It is worthwhile to note that for the proof of this theorem, it suffices to relax the definition of an (ε, α)-approximate uniform rotation X, namely, it suffices to require

only that for any $y \in S^{n-1}$, the vector Xy be an (ε, α)-approximate random sphere vector.

The proof of Theorem 6 runs more or less along the same lines as the proof of Theorem 2 in [ArM]. However, some important differences occur, and so we give the full proof below.

As in [ArM], we split the proof into two parts. First, we show that there is some constant such that, as long as $(b(\|\cdot\|)/M)$ is bounded from below by this constant, we can use one (ε_0, α)-approximate uniform operator V to construct a new norm $\|x\|_1 = \frac{\|x\|+\|Vx\|}{2}$ (with the same average M) such that $b(\|\cdot\|_1) < \lambda b(\|\cdot\|)$ for some universal constant $\lambda < 1$. In fact, we shall see that the constant λ can be chosen to be any constant greater than $1/\sqrt{2}$, while the assumed lower bound on b/M depends on the choice of this constant (and on α).

Second, we shall show that if b/M is bounded from above, then applying a finite number (this number depends on ε — the degree of approximation we wish to achieve, on the upper bound for b/M, and again on α) of approximate-uniform rotations gives a $(1+\varepsilon)$-isometric Euclidean norm. This will work exactly as in the Dvoretzky-type global theorem.

To justify the first part, we pass to the dual language and show the following.

Lemma 7. *For any $0 < \eta$ and $0 < \alpha$, there exist constants $\varepsilon_0(\eta) = \eta/3$, $C(\eta, \alpha) = \frac{30}{\sqrt{\alpha}\eta^2}$, and $c_\eta = \eta^2/8$ such that, for every dimension n and any symmetric convex body $K \subset \mathbb{R}^n$, if $C(\eta, \alpha) < \operatorname{diam}(K)/M^*(K)$, then for a random $V \in O(n)$ distributed (ε_0, α)-approximately uniformly, with probability greater than $1 - e^{-c_\eta \alpha n}$ we have*

$$\operatorname{diam}\Big(\frac{K+VK}{2}\Big) \le \frac{(1+\eta)}{\sqrt{2}} \operatorname{diam}(K).$$

Proof. First, we use Sudakov's inequality (see [Pi, p. 68]) to cover K with T Euclidean balls of radius rd, with $r < 1$ to be chosen later, where $T \le e^{2n(\frac{M^*}{rd})^2}$ (the constant 2 can actually be replaced by a constant tending to 1 as $T \to \infty$). Let x_i, $i = 1, \ldots, T$, denote the centers of this covering. We shall estimate the probability of V, distributed (ε_0, α)-approximately uniformly in $O(n)$, for which $|\langle x_i, Vx_j\rangle| \le \delta d^2$ for every $i, j = 1, \ldots, k$, where δ will be chosen later. Indeed, for fixed i and j we have

$$\mathbb{P}[V : |\langle x_i, Vx_j\rangle| \le \delta d^2] \ge \mathbb{P}[V(x_j/|x_j|) \in A_\delta],$$

where A_δ is the δ-neighborhood of

$$A = \big\{y \in S^{n-1} : \big\langle \frac{x_i}{|x_i|}, y\big\rangle \le 0\big\}$$

(observe that $\sigma(A) = 1/2$). Thus, by the assumption on the distribution of V, we see that this probability is greater than $1 - e^{-\alpha\delta^2 n}$ as long as $\delta > \varepsilon_0$.

We have at most $e^{4n(\frac{M^*}{rd})^2}$ such pairs; therefore, if, say, $e^{\alpha\delta^2 n} \ge e^{8n(\frac{M^*}{rd})^2}$, then the probability of the operators V for which this is true for all pairs (i, j) is greater than $1 - e^{-\alpha\delta^2 n/2}$. The condition amounts to $\delta \ge \sqrt{8/\alpha}(\frac{M^*}{rd})$.

Let $z \in \frac{K+VK}{2}$, that is, $z = \frac{y_1+Vy_2}{2}$, where $y_1, y_2 \in K$. Each of the points y_1, y_2 has a closest point in the net: $y_1 = x_{i_1} + rd\theta_1$ and $y_2 = x_{i_2} + rd\theta_2$, where $\theta_1, \theta_2 \in D_n$. Now

we have

$$\begin{aligned}\Big|\frac{y_1+Vy_2}{2}\Big|^2 &= \Big|\frac{x_{i_1}+rd\theta_1+Vx_{i_2}+rdV\theta_2}{2}\Big|^2\\ &= \frac{|x_{i_1}|^2}{4}+\frac{|x_{i_2}|^2}{4}+\frac{\langle x_{i_1},Vx_{i_2}\rangle}{2}\\ &\quad+\frac{|rd(\theta_1+V\theta_2)|^2}{4}+\frac{\langle x_{i_1}+Vx_{i_2},rd(\theta_1+V\theta_2)\rangle}{2}\\ &\le d^2\Big(\frac{1}{2}+\frac{\delta}{2}+r^2+2r\Big).\end{aligned}$$

Choosing, for example, $r=\eta/5$, $\delta=\eta/2$, we get the inequality

$$\Big|\frac{y_1+Uy_2}{2}\Big|\le d\sqrt{\frac{1}{2}+\eta}\le d\frac{1+\eta}{\sqrt{2}},$$

and the condition above takes the form $\frac{d}{M^*}\ge\frac{\sqrt{8}10}{\sqrt{\alpha}\eta^2}$. Thus, for any fixed α,η we may take $C(\eta,\alpha)=30/(\sqrt{\alpha}\eta^2)$, $c_\eta=\eta^2/8$, and $\varepsilon_0=\eta/3$. □

To complete the proof of Theorem 6 we need the following proposition, which is in fact the usual global Dvoretzky-type theorem, that is, Theorem 1, for the more general approximate-uniform operators, in the case where b/M is bounded from above.

Theorem 8. *There exists a universal c' such that, for any $0<\alpha$, C, and $c'C/\sqrt{\alpha n}<\varepsilon<1$, there exist constants $\varepsilon_0(\varepsilon,C)$, $c_1(\varepsilon)$, and $C_1(C,\varepsilon,\alpha)$ such that, for any dimension n and any $K\subset\mathbb{R}^n$ satisfying $\frac{b(K)}{M(K)}<C$, with probability greater than $1-e^{-c_1 n}$ the random orthogonal (ε_0,α)-approximate uniform transformations $V_1,\dots,V_{C_1}\in O(n)$ satisfy*

$$(1-\varepsilon)M|x|\le\frac{1}{C_1}\sum_{i=1}^{C_1}\|V_ix\|_K\le(1+\varepsilon)M|x| \tag{3}$$

for every $x\in\mathbb{R}^n$. We may take $\varepsilon_0=\varepsilon/(13C)$, $C_1=C_0C^2\ln(1+1/\varepsilon)/(\varepsilon^2\alpha)$, and $c_1=\ln(11/\varepsilon)$ (for a universal C_0).

Proof. We shall use a Bernstein-type inequality in the form presented in [Ar]. Notice that, since we are dealing with approximate-uniform random variables, we cannot use the standard Bernstein inequality (see, e.g., [BLM]) in the straightforward way, as is usually done in the proofs of Dvoretzky-type theorems, because we have no control on the near-zero parts of the random variable $\|V_ix\|-M$.

We define $Y_i=\|V_ix\|-M_1$, where M_1 is the Levy mean of the function $\|x\|$ on S^{n-1}, so that $\sigma(\{x:\|x\|<M_1\})=1/2$. Note that $M_1<M<M_1+c''b/\sqrt{n}$ for a universal c'', see, e.g., [MS1] again. Thus, the assumption that b/M is bounded from above implies that the ratio between M_1 and M is extremely close to 1. By standard concentration (see, e.g., [MS1]) and by the definition of approximate uniformity, as long as $\delta>\varepsilon_0(b/M)$, we have

$$\mathbb{P}[|Y_i|>\delta M]\le 2e^{-\alpha\delta^2(M/b)^2n}.$$

Now we use Theorem 1 of [Ar]. In the notation of that theorem, with $p=2$ and $\delta M=t$, we let $K_0=\varepsilon_0 b$ and, assuming that $\delta>(b/M)\sqrt{2\ln 2}/\sqrt{\alpha n}$, we may take $K=\alpha n/(2b^2)$. Then [Ar, Theorem 1] implies that for some universal c_0, for $\delta>\max(3\varepsilon_0(b/M),3\sqrt{2}(1+\ln 2)(b/M)/\sqrt{\alpha n})$ we have

$$\mathbb{P}\Big[\frac{1}{N}\sum_{i=1}^{N}|Y_i|>\delta M\Big]\le c_0e^{-N\delta^2(M/b)^2\alpha n/18}.$$

The triangle inequality shows that, under the same conditions on δ,

$$\mathbb{P}\Big[M_1 - \delta M \le \frac{1}{N}\sum_{i=1}^{N}\|V_i x\| \le M_1 + \delta M\Big] \ge 1 - c_0 e^{-N\delta^2 (M/b)^2 \alpha n/18}.$$

Recall that $M < M_1(1 + c''C/\sqrt{n})$, so that $M < 2M_1$ for n sufficiently large. Thus, we may take, say, $\delta = \varepsilon/4$. To make sure this choice of δ works well with the condition in the theorem we have used, we must have $\varepsilon > 12\varepsilon_0 C$ and $\varepsilon > 12\sqrt{2}(1+\ln 2)(b/M)/\sqrt{\alpha n}$. The second condition is satisfied for n large enough, namely, $n \ge c_0 C^2/\alpha\varepsilon^2$. The first is ensured by taking, e.g., $\varepsilon_0 = \varepsilon/(13C)$.

Therefore, for a specific x we have

$$\mathbb{P}\Big[(1-\varepsilon/2)M_1 \le \frac{1}{N}\sum_{i=1}^{N}\|V_i x\| \le (1+\varepsilon/2)M_1\Big] \ge 1 - c_0 e^{-N\varepsilon^2\alpha n/(144C^2)}.$$

However, we need this to be valid for an entire $\varepsilon/5$-net on the sphere (say), in order to use successive approximation and get a result for the whole sphere. This net has cardinality at most $(11/\varepsilon)^n$, so we need to choose N with, say,

$$c_0 e^{N\varepsilon^2\alpha/(144C^2)} > (11/\varepsilon)^2,$$

which amounts to $N = C_1 = C^2\ln(121/c_0\varepsilon^2)/(\varepsilon^2\alpha)$. Then the event in question happens with probability greater than $1 - (11/\varepsilon)^{-n}$. We get the upper bound $M_1(1+\varepsilon/2)/(1-\varepsilon/5) < (1+7\varepsilon/8)M_1$ for all points on the sphere, and the lower bound $M_1(1-9\varepsilon/10)$. Since the ratio between M and M_1 is between 1 and $(1+c''C/\sqrt{n})$ under the assumptions of the theorem, we may (for n sufficiently large compared to ε, as is assumed) replace M_1 by M and get the conclusion of the theorem. □

Proof of Theorem 6. Let $\alpha > 0$, and let $c_0/(\alpha\sqrt{n}) < \varepsilon < 1$ (with a universal c_0 to be chosen later in the proof). First, we use Lemma 7 with $\eta = 0.1$ say, and apply it $N_1 = C_2 + C_3\log(b/M)$ times, where $C_3 = 1/\ln(\sqrt{2}/1.1)$ and $C_2 = C_3\ln(3000/\sqrt{\alpha})$. We see that, for the N_1 operators $V_1, \dots, V_{N_1}$, with probability greater than $1 - N_1 e^{-c\alpha n}$, for the norm defined by

$$\|x\|_{K_0} = \frac{1}{2^{N_1}}\sum_{\delta\in\{0,1\}^{N_1}}\Big\|\prod_{i=1}^{N_1} V_i^{\delta_i} x\Big\|_K,$$

we have $b(\|\cdot\|_{K_0})/M(|\cdot\|_{K_0}) \le 3000/\sqrt{\alpha}$.

Next, we apply Theorem 8 with $C = 3000/\sqrt{\alpha}$, so that the assumption on ε is satisfied provided c_0 is chosen sufficiently large. As a result, the C_1 operators, for the choice of C_1 as in the statement of the theorem, satisfy the desired inequality with probability greater than $1 - (\frac{\varepsilon}{11})^n$. The final probability, $1 - N_1 e^{-c\alpha n} - (\frac{\varepsilon}{11})^n$, is in its turn greater than $1 - e^{-n\min(\alpha, \ln(11/\varepsilon))/C_0}$ for some universal C_0, because n is assumed large. □

3.4. Concentration. We return to the main line of this paper, namely, to proving Theorems 3 and 4.

To be able to use the fact that for any $0 \ne x$ the vectors $U_0 x$, $U_1 x$, $U_2 x$, ... are mutually orthogonal, we need to recall the strong concentration result that is true for orthogonal frames. We recall the basic facts about concentration on the Stiefel manifold, which will also be used in §4.

The Stiefel manifold $W_{n,k} = W_k(\mathbb{R}^n)$ consists of systems of k orthonormal vectors in $\mathbb{R}^n$ with the metric

$$d\left((e_1,\dots,e_k),(f_1,\dots,f_k)\right) = \left(\sum_{i=1}^k |e_i - f_i|^2\right)^{1/2}.$$

Being a homogeneous space of the unitary group, the Stiefel manifold possesses a strong concentration property (see [GrM])

Proposition 9 (Concentration on Stiefel)**.** *For any function $f : W_{n,k} \to \mathbb{R}$ with Lipschitz constant L, for some number M_f (the Levy mean of f) we have*

$$\mathbb{P}[U \in O(n) : |f(Ue_1,\dots,Ue_k) - M_f| > \varepsilon] \le e^{-c' n\varepsilon^2/L^2}$$

with a universal c'. Here e_1, …, e_k are the first k elements of the standard unit vector basis.

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ with mean M and maximum b. Consider the function $f : W_{n,k} \to \mathbb{R}$ defined by $f(x_1,\dots,x_k) = \frac{1}{k}\sum_{i=1}^k \|x_i\|$. It is Lipschitz with constant $b/\sqrt{k}$, because

$$\begin{aligned}|f(x_1,\dots,x_k) - f(y_1,\dots,y_k)| &\le \frac{1}{k}\sum_{i=1}^k \|x_i - y_i\| \\ &\le \frac{b}{k}\sum_{i=1}^k |x_i - y_i| \le \frac{b}{\sqrt{k}} d\left((x_1,\dots,x_k),(y_1,\dots,y_k)\right).\end{aligned}$$

Therefore, the above proposition implies that for any orthonormal k-tuple $(x_1,\dots,x_k)$ and any $0 < \varepsilon < 1$ we have

$$\mathbb{P}\Big[U : M_0 - \varepsilon M \le \frac{1}{k}\sum_{i=1}^k \|Ux_i\| \le M_0 + \varepsilon M\Big] \ge 1 - e^{-c'\varepsilon^2 nk(M/b)^2}, \tag{4}$$

where c' is universal.

Here we use the letter M_0 to denote the Levy mean of the function $f(U) = \frac{1}{k}\sum\limits_{i=1}^k \|Ue_i\|$. The Lipschitz constant of this function is at most b. The usual mean (average) of this function is M, and standard arguments (see, e.g., [MS1]) show that the difference between M_0/M and 1 is very small.

3.5. Bodies with logarithmically small b/M. Now we prove Theorem 3. Recall that the conditions of the theorem say that for the body $K \subset \mathbb{R}^n$ we have $C_0(b(K)/M(K))^2 \le \log n$, where C_0 depends only on the parameter $0 < \varepsilon < 1$ chosen in advance. We have denoted $k = [C_0(b/M)^2]$, and the operators U_0, …, U_{k-1} were defined in §2.

Proof of Theorem 3. Let $\mathcal{N} = \{z^{(j)}\}_{j=1}^T$ be a δ-net on the sphere S^{n-1}, so that we can have $T \le e^{n\ln(3/\delta)}$. We fix $\delta = \varepsilon/(2+2\varepsilon)$ (note that $\varepsilon/5 < \delta < \varepsilon/2$). Let C_0 be a constant depending only on ε, to be determined later. Denote $k = C_0(b/M)^2$ (we ignore taking the integral part for notational convenience; this will not change anything except possibly the constants), and we have assumed that $k \le \log n$. Consider the operators U_0, …, U_{k-1} defined in §2. For every $z^{(j)}$, we use these operators to create k vectors $z_i^{(j)} = U_i z^{(j)}$, which are mutually orthogonal. Now we use estimate (4), which promises that, for a random $v \in O(n)$ and any fixed j,

$$\mathbb{P}\Big[v : M_0 - \frac{\varepsilon}{2}M \le \frac{1}{k}\sum_{i=0}^{k-1} \|Vz_i^{(j)}\| \le M_0 + \frac{\varepsilon}{2}M\Big] \ge 1 - e^{-c' C_0 \varepsilon^2 n/4}.$$

To simplify the presentation, we shall use the easily verified fact that $M < M_0 + c'' \frac{b}{\sqrt{n}}$ for some universal c'' (see, e.g., [MS1]); under our assumptions, when n is not too small, this implies that for some other universal c_1' we have

$$\mathbb{P}\Big[V : \Big(1 - \frac{\varepsilon}{2}\Big) M_0 \le \frac{1}{k} \sum_{i=0}^{k-1} \|V z_i^{(j)}\| \le \Big(1 + \frac{\varepsilon}{2}\Big) M_0\Big] \ge 1 - e^{-c_1' C_0 \varepsilon^2 n}.$$

If $C_0 = 2\ln(3/\delta)/(c_1'\varepsilon^2)$, then with probability greater than $1 - e^{-n\ln(3/\delta)}$, for every $z^{(j)} \in \mathcal{N}$ we have

$$(1 - \varepsilon/2) M_0 \le \frac{1}{k} \sum_{i=0}^{k-1} \|V U_i z^{(j)}\| \le (1 + \varepsilon/2) M_0. \tag{5}$$

By successive approximation, every $z \in S^{n-1}$ can be written as

$$z = z^{(j_1)} + \theta_1 z^{(j_2)} + \theta_2 z^{(j_3)} + \cdots$$

where $z^{(j_i)} \in \mathcal{N}$ and $|\theta_i| \le \delta^i$. Therefore, using the triangle inequality and the bound on each vector in $\mathcal{N}$, we see that for the same operators V satisfying (5) and for any $z \in S^{n-1}$ we also have

$$\frac{1}{k} \sum_{i=0}^{k-1} \|V U_i z\| \le \frac{(1 + \varepsilon/2)}{(1 - \delta)} M_0 = (1 + \varepsilon) M_0. \tag{6}$$

On the other hand, the lower bound in (5), the upper bound in (6), the triangle inequality, and the fact that any $z \in S^{n-1}$ can be written as $z = z^{(j)} - \theta y$ with $z^{(j)} \in \mathcal{N}$, $y \in S^{n-1}$, and $|\theta| < \delta$, imply that

$$\frac{1}{k} \sum_{i=0}^{k-1} \|V U_i z\| \ge (1 - \varepsilon/2) M_0 - \delta(1 + \varepsilon) M_0 = (1 - \varepsilon) M_0$$

for any $z \in S^{n-1}$. □

3.6. General convex bodies. In the proof of Theorem 4 we shall use the following lemma.

Lemma 10. *For any constant $C_0 > 1$, there exists a constant $C_2 > 1$ such that for any dimension $n = 2^k$, for some integer k, and any norm $\|\cdot\|$ defined on $\mathbb{R}^n$ with average M and maximum b, if $C_0(b/M)^2 > k$, then the probability of choosing uniformly a random rotation V such that the rotations $U_0, \ldots, U_{k-1}$ defined above satisfy*

$$\frac{1}{k} \sum_{i=0}^{k-1} \|V U_i x\|_K \le C_2 \frac{b}{\sqrt{k}} |x|$$

for every $x \in \mathbb{R}^n$, is greater than $1 - 5^{-n}$.

Proof of Lemma 10. The proof of this lemma is quite similar to that of the upper estimate in Theorem 3. We consider a 1/2-net $\mathcal{N} = \{z^{(j)}\}_{j=1}^T$ on the sphere S^{n-1}, so that we can have $T \le 5^n$.

Consider the operators $U_0, \ldots, U_{k-1}$ defined in §2. For every $z^{(j)}$ we use these operators to create k vectors $z_i^{(j)} = U_i z^{(j)}$, which are mutually orthogonal. Now we use the estimate (4), which is valid also for $\varepsilon = t > 1$ (then only one side of the inequality makes sense). This estimate promises that, for a random $V \in O(n)$ and any fixed j,

$$\mathbb{P}\Big[V : \frac{1}{k} \sum_{i=0}^{k-1} \|V z_i^{(j)}\| \le tb/\sqrt{k}\Big] \ge 1 - e^{-c' n t^2/2},$$

provided $(tb/M)/\sqrt{k} > 2$. (We use the fact that $M > M_0$.) We choose

$$t = \max(\sqrt{4 \ln 5/c'}, 2C_0^{1/2}) =: C_2/2,$$

so that the above condition is satisfied, by the assumptions of the lemma. Then, with probability $1 - 5^{-n}$, for every point in the net we have

$$\frac{1}{k} \sum_{i=0}^{k-1} \|VU_i z^{(j)}\| \le (C_2/2) b/\sqrt{k}. \tag{7}$$

Successive approximation yields the bound $C_2 b/\sqrt{k}$ for all points on the sphere, which completes the proof of the lemma by homogeneity. □

To prove Theorem 4, we shall iterate the above lemma, with the choice of C_0 coming from Theorem 3, until we are under the conditions of Theorem 3.

Proof of Theorem 4. We start with a general norm and denote $\|\cdot\|_1 := \|\cdot\|$ and $b_1 := b(\|\cdot\|)$. If $C_0(b/M)^2 \le \log n$, then one random rotation suffices to construct explicitly $C_0(b/M)^2$ rotations that together satisfy (2). If not, then from Lemma 10 we use one random rotation V to build $\log n$ operators VU_i, which determine for us a new norm

$$\|\cdot\|_2 := \frac{1}{k} \sum_{i=0}^{k-1} \|VU_i z^{(j)}\|_1.$$

Clearly, this norm has the same average M as the original norm, and by the lemma, with probability greater than $1 - 5^{-n}$ we have $b_2 = b(\|\cdot\|_2) \le C_2 b_1/\sqrt{\log n}$, where C_2 is universal (because it depends only on C_0, which was universal).

Repeating this iterative process, for $j = 2, 3, \ldots$ we create the norms

$$\|\cdot\|_{j+1} := \frac{1}{k} \sum_{i=0}^{k-1} \|V_j U_i z^{(j)}\|_j,$$

which all have the same average M and satisfy

$$b_{j+1} \le C_2 b_j/\sqrt{\log n} \le b_1 (C_2/\sqrt{\log n})^j.$$

We may continue this process until the first j for which $C_0(b_j/M)^2 \le \log n$. The above implies that this must happen after at most j_0 steps with

$$j_0 = \frac{\log\left(C_0 \frac{(b/M)^2}{\log n}\right)}{\log\left(\frac{\log n}{C_2^2}\right)} + 1.$$

In creating the $(j+1)$th norm from the jth norm, we use up exactly one random operator, which we combine with the U_i's to create the next norm. Therefore, we have used at most $(j_0 - 1)$ random operators in this process. Now we are back in the position of Theorem 3, and use only one more random operator to complete the theorem. The probability estimate is $1 - j_0 5^{-n} - e^{-c_1(\varepsilon)n}$, where $c_1(\varepsilon)$ comes from Theorem 3. The final operators W_k are the products of terms of the form $V_j U_i$, that is, $\prod_j V_j U_{i_j}$. □

3.7. Eliminating the condition $n = 2^k$. Since the operators defined in §2 are in $O(n)$ with $n = 2^k$ for some integer k, we were forced to restrict ourselves to convex bodies in $\mathbb{R}^n$ for these special n. However, any $\mathbb{R}^n$ can be regarded as a subspace of R^{2^k} with $k \le 2 \log n$. Therefore, any convex body can be assumed to lie in $\mathbb{R}^{2^k}$, and in such a way that its diameter stays the same. It can be shown easily that its mean width M^* is changed by at most $1/\sqrt{2}$: it does not remain the same because now integration is over a larger sphere, but the volume distribution on this sphere, since its dimension is at most

twice the dimension of the subsphere, is concentrated at a distance of at most $1/\sqrt{2}$ from the subsphere.

More precisely, when working in the dual setting and considering norms, we shall define a (semi)norm of a point in $\mathbb{R}^{2^k}$ as the norm of the projection of this point to the first n coordinates. This is a seminorm and not a norm, but it is easily seen that all the above results are valid also for seminorms. The maximum b of the new seminorm is the same as the maximum of the original one, and, since $n > 2^{k-1}$, the average M is changed by a factor of at most $1/\sqrt{2}$. This is standard and can be seen by passing to Gaussian averages; see, e.g., [GiM].

Once the body is embedded in R^{2^k}, we may operate on it with the k operators defined in §2, and arrive at the same results as in the preceding subsections.

§4. An application to zigzag bodies

In the paper [AFM1], following similar constructions by Barron and Cheang (see [BC]), it was shown that, picking randomly $N = Cn\ln(1/\varepsilon)/\varepsilon^2$ points distributed uniformly on the unit sphere and constructing the so-called "zigzag" body

$$\mathcal{K} = \Big\{x \in \mathbb{R}^n : \exists j_1, \dots, j_{[N/2]} \text{ with } |\langle x, z_{j_k}\rangle| < \frac{c_0}{\sqrt{n}}\Big\},$$

where c_0 is such that $c_0/\sqrt{n}$ is the Levy mean of the function $|\langle x, \theta\rangle|$ on the sphere, with high probability one gets a body (usually not convex) that is ε-close to the Euclidean ball D_n. The advantage of this approximating body is that it is defined by using relatively few inequalities all of which are linear. For history and more details see [AFM1] and the references therein.

In this section we show how another type of a "zigzag" body that approximates the Euclidean ball can be constructed with less random selections, with the help of the operators defined in §2. Using these operators, we replace a logarithmic proportion of the random points needed for the construction by explicit steps.

Our new definition of a zigzag body is as follows: we pick N points $z_1, \dots, z_N$ of the sphere S^{n-1} independently and uniformly on S^{n-1}, and introduce the (not necessarily convex) body

$$\mathcal{K} = \Big\{x \in \mathbb{R}^n : \exists j_1, \dots, j_{[N/2]} \text{ such that } \frac{1}{\log n}\sum_{i=0}^{\log n - 1} |\langle x, U_i z_{j_k}\rangle| < \frac{c_1}{\sqrt{n}}\Big\} \tag{8}$$

where c_1 is defined so that the Levy mean of the function $\frac{1}{\log n}\sum_{i=1}^{\log n} |\langle x, e_i\rangle|$ on the sphere equals $c_1/\sqrt{n}$. (This c_1 does depend on n, but converges to a universal constant as n tends to infinity, as did c_0 defined above.)

We prove that, with this new definition of $\mathcal{K}$, the following is true.

Theorem 11. *There exist universal constants C, c such that, for every dimension $n = 2^k$ for some integer k, and every $\frac{c}{\sqrt{\log n}} < \varepsilon < 1$, if $N = [C(n/\log n)\ln(1/\varepsilon)/\varepsilon^2]$ and $z_1, \dots, z_N$ are independent random points distributed uniformly on S^{n-1}, then, with probability greater than $1 - e^{-cn}$, the set $\mathcal{K}$ defined in (8) satisfies*

$$(1-\varepsilon)D_n \subset \mathcal{K} \subset (1+\varepsilon)D_n.$$

Remark. The constraint $n = 2^k$ here is not important, because we may always intersect the zigzag body in $\mathbb{R}^{2^k}$ with the subspace spanned by the first $n > 2^{k-1}$ unit vectors. The zigzag body obtained in $\mathbb{R}^n$ in this way is defined by using the same number of vectors (which are now the projections of the vectors defined above). If the original body is $(1+\varepsilon)$-isometric to a Euclidean ball, then so will be its intersection with the subspace.

The proof of Theorem 11 is similar to the proofs of the zigzag theorems in [AFM1], which employ a method that appeared in the paper [MP]. However, the probabilities involved are slightly different, because the usage of the operators U_i leads to amplification of probability (as is apparent from (4)). We begin with a simple and useful lemma.

Lemma 12. *There exists some universal constant $C > 1$ such that if $N \geq Cn/\log n$, then for $z_1, \ldots, z_N$ independent random points distributed uniformly on S^{n-1}, we have*

$$\mathbb{P}\Big[\forall y, \frac{1}{N\log n}\sum_{j=1}^{N}\sum_{i=0}^{\log n-1} |\langle U_i z_j, y\rangle| \leq C|y|/\sqrt{n}\Big] \geq 1 - e^{-N\log n}. \tag{9}$$

Proof of Lemma 12. We shall use the upper bound in the concentration estimate (4) for the seminorm $\|z\| = |\langle z, y\rangle|$ with a fixed y in some 1/2-net of the sphere. So, for $t > 2$,

$$\mathbb{P}\Big[z : \frac{1}{\log n}\sum_{i=1}^{\log n} \|U_i z\| \leq t\frac{c_1}{\sqrt{n}}\Big] \geq 1 - e^{-c_2 t^2 \log n}$$

where c_2 is universal. Now we need a Bernstein-type argument, with this tail estimate. For example, we may use Theorem 1 in [Ar], which implies that

$$\mathbb{P}\Big[z_1, \ldots, z_N : \frac{1}{N}\sum_{j=1}^{N}\frac{1}{\log n}\sum_{i=0}^{\log n-1} \|U_i z_j\| \leq t\frac{c_1}{\sqrt{n}}\Big] \geq 1 - e^{-c_3 t^2 N\log n}$$

for every $t > 6$. Therefore, as long as $N \geq c_4 n/\log n$ for some fixed c_4 (the constant c_4 can in fact be fixed as a small constant, because t can compensate for its smallness), we may choose t so large that (9) can be achieved for an entire net of y's of cardinality at most 5^n, and by successive approximation we arrive at the conclusion of Lemma 9 for the whole sphere, with $C = 2tc_1$. □

Proof of Theorem 11. We start with showing that $D \subset (1+\varepsilon)D_n$. We cover $(1+\varepsilon)S^{n-1}$ by balls of radius $\alpha\varepsilon$ with α to be determined later. As is well known, there exists such a covering of cardinality less than $e^{\ln(3/\alpha\varepsilon)n}$. Consider one of these balls, centered at x_0. By (4), we know that

$$\mathbb{P}\Big[z \in S^{n-1} : \frac{1}{\log n}\sum_{i=1}^{\log n} |\langle U_i z, x_0\rangle| < \frac{c_1}{\sqrt{n}}(1+\varepsilon/2)\Big] \leq e^{-c_5\varepsilon^2\log n} \tag{10}$$

for some universal c_5 (because $|x_0| = (1+\varepsilon)$).

Therefore, there is a high probability that more than $(3/4)N$ of the indices j will satisfy

$$\frac{1}{\log n}\sum_{i=1}^{\log n} |\langle x, U_i z_j\rangle| > (1+\varepsilon/2)\frac{c_1}{\sqrt{n}}.$$

More precisely, the Chernoff bounds (see, e.g., [AFM2]) with $p = e^{-c_5\varepsilon^2\log n}$ and $\beta = 1/4$ imply that

$$\begin{aligned}\mathbb{P}\Big[&z_1, \ldots, z_N \in S^{n-1} : \\ &\text{for at least } (3/4)N \text{ of the indices } j \\ &\text{we have } \frac{1}{\log n}\sum_{i=1}^{\log n} |\langle x, U_i z_j\rangle| \geq (1+\varepsilon/2)\frac{c_1}{\sqrt{n}}\Big] \\ &\geq 1 - e^{-Nc_6\varepsilon^2\log n}.\end{aligned}$$

Notice that this is only true for n sufficiently large compared to ε, namely, for $n > e^{(C/\varepsilon)^2}$. This explains the condition on ε in the statement of the theorem. There is a way to avoid the assumption on ε in the theorem and arrive at a slightly more complicated expression valid for all $\varepsilon > 0$. For this, we can change $3/4$ to a smaller number, close to $1/2$, depending on ε, and also modify the bound (10) to be meaningful for all ε. We shall not pursue this here; for a similar analysis see [AFM1].

We see that, for the above to happen with high probability for the entire net simultaneously, N must be of the order $(\log(c/\varepsilon\alpha)/\varepsilon^2)n/\log n$ (later α will be universal).

If $x \in \mathcal{K}$, then for half of the indices j we have

$$\frac{1}{\log n} \sum_{i=0}^{\log n - 1} |\langle x, U_i z_j \rangle| < \frac{c_1}{\sqrt{n}}.$$

In particular, for any fixed x_0, there are at least $N/4$ indices j for which, at the same time, also

$$\frac{1}{\log n} \sum_{i=0}^{\log n - 1} |\langle x_0, U_i z_j \rangle| \geq (1 + \varepsilon/2) \frac{c_1}{\sqrt{n}}.$$

Thus,

$$\frac{1}{\log n} \sum_{i=0}^{\log n - 1} |\langle x_0 - x, U_i z_j \rangle| \geq (\varepsilon/2) \frac{c_1}{\sqrt{n}}$$

for each of these indices j. Together with Lemma 9, this implies the estimate

$$|x - x_0| \geq \frac{c_1}{C} \frac{1}{4} (\varepsilon/2).$$

Thus, as long as $\alpha < \frac{c_1}{8C}$, we have

$$\mathcal{K} \cap B(x_0, \alpha\varepsilon) = \varnothing,$$

and this is true for every x_0 in the net. Therefore, $\mathcal{K} \subset (1 + \varepsilon) D_n$.

The second inclusion in the statement of the theorem admits much the same proof, which is thus omitted. □

References

[AS] N. Alon and J. H. Spencer, *The probabilistic method*, Wiley-Intersci. Ser. in Discrete Math. Optim., Wiley–Intersci., New York, 2000. MR1885388 (2003f:60003)

[Ar] S. Artstein-Avidan, *A Bernstein–Chernoff deviation inequality, and geometric properties of random families of operators*, Israel J. Math. **156** (2006), 187–204. MR2282375

[ArM] S. Artstein-Avidan and V. Milman, *Logarithmic reduction of the level of randomness in some probabilistic geometric constructions*, J. Funct. Anal. **235** (2006), 297–329. MR2216448 (2007a:46009)

[AFM1] S. Artstein-Avidan, O. Friedland, and V. Milman, *Geometric applications of Chernoff-type estimates and a zigzag approximation for balls*, Proc. Amer. Math. Soc. **134** (2006), 1735–1742. MR2204286 (2006k:46014)

[AFM2] ———, *More geometric applications of Chernoff inequality*, Geometric Aspects of Functional Analysis, Springer Lecture Notes in Math. (to appear).

[BC] A. Barron and G. Cheang, *A better approximation for balls*, J. Approx. Theory **104** (2000), no. 2, 183–203. MR1761898 (2001d:41018)

[BN] A. Ben-Tal and A. Nemirovski, *On polyhedral approximations of the second-order cone*, Math. Oper. Res. **26** (2001), no. 2, 193–205. MR1895823 (2003d:90119)

[BLM] J. Bourgain, J. Lindenstrauss, and V. Milman, *Minkowski sums and symmetrizations*, Geometric Aspects of Functional Analysis (1986/87), Lecture Notes in Math., vol. 1317, Springer, Berlin, 1988, pp. 44–66. MR0950975 (89g:46025)

[GiM] A. Giannopoulos and V. Milman, *Euclidean structure in finite dimensional normed spaces*, Handbook of the Geometry of Banach Spaces, Vol. I, North-Holland, Amsterdam, 2001, pp. 707–779. MR1863705 (2003b:46008)

[GrM] M. Gromov and V. Milman, *A topological application of the isoperimetric inequality*, Amer. J. Math. **105** (1983), no. 4, 843–854. MR0708367 (84k:28012)

[M] V. Milman, *The concentration phenomenon and linear structure of finite-dimensional normed spaces*, Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Berkeley, CA, 1986), Amer. Math. Soc., Providence, RI, 1987, pp. 961–975. MR0934298 (89h:46029)

[MP] V. Milman and A. Pajor, *Regularization of star bodies by random hyperplane cut off*, Studia Math. **159** (2003), no. 2, 247–261. MR2052221 (2005e:52009)

[MS1] V. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces* (*with an appendix by M. Gromov*), Lecture Notes in Math., vol. 1200, Springer-Verlag, Berlin, 1986. MR0856576 (87m:46038)

[MS2] ———, *Global versus local asymptotic theories of finite-dimensional normed spaces*, Duke Math. J. **90** (1997), no. 1, 73–93. MR1478544 (98m:46005)

[Pi] G. Pisier, *The volume of convex bodies and Banach space geometry*, Cambridge Tracts in Math., vol. 94, Cambridge Univ. Press, Cambridge, 1989. MR1036275 (91d:52005)

School of Mathematical Science, Tel Aviv University, Ramat Aviv, 69978, Tel Aviv, Israel
E-mail address: shiri@post.tau.ac.il

School of Mathematical Science, Tel Aviv University, Ramat Aviv, 69978, Tel Aviv, Israel
E-mail address: milman@post.tau.ac.il

Received 1/AUG/2006
Originally published in English