PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 134, Number 6, Pages 1735–1742
S 0002-9939(05)08144-X
Article electronically published on December 14, 2005

GEOMETRIC APPLICATIONS OF CHERNOFF-TYPE ESTIMATES AND A ZIGZAG APPROXIMATION FOR BALLS

S. ARTSTEIN-AVIDAN, O. FRIEDLAND, AND V. MILMAN

(Communicated by N. Tomczak-Jaegermann)

Abstract. In this paper we show that the euclidean ball of radius 1 in $\mathbb{R}^n$ can be approximated up to $\varepsilon > 0$, in the Hausdorff distance, by a set defined by $N = C(\varepsilon)n$ linear inequalities. We call this set a ZigZag set, and it is defined to be all points in space satisfying 50% or more of the inequalities. The constant we get is $C(\varepsilon) = C\ln(1/\varepsilon)/\varepsilon^2$, where C is some universal constant. This should be compared with the result of Barron and Cheang (2000), who obtained $N = Cn^2/\varepsilon^2$. The main ingredient in our proof is the use of Chernoff's inequality in a geometric context. After proving the theorem, we describe several other results which can be obtained using similar methods.

The aim of this paper is to demonstrate how the well-known Chernoff estimates from probability theory can be used in a geometric context for a very broad spectrum of problems, and how they lead to new and improved results. We will briefly describe Chernoff bounds, and then present the motivation for, and the proof of, the following theorem:

Theorem 1. *There exist universal constants C, c such that for every dimension n, and every $0 < \varepsilon < 1$, letting $N = [Cn\ln(1/\varepsilon)/\varepsilon^2]$, if $z_1, \ldots, z_N$ are random points with respect to Lebesgue measure σ on the sphere S^{n-1}, then with probability greater than $1 - e^{-cn}$, the set*

$$\mathcal{K} = \{x \in \mathbb{R}^n : \exists i_1, \ldots, i_{[N/2]} \text{ with } |\langle x, z_{i_j}\rangle| < \frac{c_0}{\sqrt{n}}\}$$

satisfies

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

where c_0 denotes the constant (depending on n, but converging to a universal constant as $n \to \infty$) for which $\sigma(u \in S^{n-1} : |\langle \theta, u\rangle| \le \frac{c_0}{\sqrt{n}}) = 1/2$ for some $\theta \in S^{n-1}$.

This theorem improves on a result of Barron and Cheang [CB], as we discuss further on in this note.

After proving this theorem, we will state several other theorems which can be proved using similar methods, and whose proofs, along with more theorems, will appear in [AFM]. This probabilistic approach was already used in our field in the paper [MP], a paper which is closely related to the proof of Theorem 1.

Received by the editors October 26, 2004 and, in revised form, January 18, 2005.
2000 *Mathematics Subject Classification.* Primary 46B07; Secondary 60D05, 46B09.
This research was partially supported by BSF grant 2002-006 and by FP6 Marie Curie Actions "PHD".

We will, throughout, omit the notation $[\cdot]$ of integer values, and assume the numbers we deal with are integers when needed, to avoid notational inconvenience.

The probabilistic approach

The following lemma, which is a version of Chernoff's inequalities, gives an estimate for the probability that at least βN trials out of N succeed, when the probability of a success in one trial is at least p (see e.g. [HR]).

Lemma 2. *Let Z_i be independent Bernoulli random variables with mean $0 < p < 1$ (that is, Z_i takes value 1 with probability p and value 0 with probability $(1-p)$). Then for every $\beta < p$ we have*

$$\mathbb{P}[Z_1 + \cdots + Z_N \geq \beta N] \geq 1 - e^{-NI(\beta,p)},$$

and for every $\beta > p$ we have

$$\mathbb{P}[Z_1 + \cdots + Z_N > \beta N] \leq e^{-NI(\beta,p)},$$

where $I(\beta,p) = \beta \ln \frac{\beta}{p} + (1-\beta)\ln\frac{1-\beta}{1-p}$.

A ZigZag approximation for balls

We address the question of approximating the euclidean ball by a simpler set. In many contexts, polytopes are considered to be the simplest sets available, being the intersection of some number of half-spaces, or in other words the set of all points satisfying some list of N linear inequalities. However, it is well known and easy to check that to construct a polytope which is ε-close, in the Hausdorff metric, to the euclidean ball $D_n \subset \mathbb{R}^n$, one needs to use exponentially many half-spaces, $N \geq e^{Cn\ln(1/\varepsilon)}$. (This can be seen by assuming the polytope is inscribed in D_n, and estimating from above the volume of the cap that each half-space cuts off the sphere S^{n-1}.) This is a huge number, and so a different kind of approximation was suggested, first used by Cybenko, and by Hornik, Stinchcombe, and White; see [C], [HSW].

The first good bounds in such an approximation result (we describe the approximating set below) were given by Barron; see [B]. These sets are implemented by what is called single hidden layer neural nets or perception nets, and we will use the simplest version of such sets, for which we suggest the name "ZigZag approximation".

The approximating set is the following: It is no longer convex, but is still described by a list of linear inequalities. Given a set of N inequalities, and a number $k \leq N$, the set consists of all points satisfying no less than k of the N inequalities. We learned of this approximation from a paper by Barron and Cheang [CB], where they showed that there exists a universal constant C such that for any dimension n, one can find $N = C(n/\varepsilon)^2$ linear inequalities, such that the set of points satisfying at least k of the N inequalities is ε-close, in the Hausdorff metric, to D_n (where k is some proportion of N). This is already a huge improvement, from a set described by an exponential number of inequalities to a polynomial number.

In this section we improve their estimate to $N = Cn\ln(1/\varepsilon)/\varepsilon^2$ linear inequalities, and we use $k = N/2$. The formulation of our result is given in Theorem 1.

To make the exposition clearer, we will first prove a different theorem, Theorem 3, which gives a weaker estimate, but is still in some cases a considerable improvement of the result from [CB], and in the case of ε not too small gives that $N = Cn\ln n/\varepsilon^2$ linear inequalities are enough. This result has a rather straightforward proof. We will then employ more involved techniques from [MP] to prove the stronger version.

As before, denote by c_0 the constant (depending on n, but converging to a universal constant as $n \to \infty$) for which for some (or any) $\theta \in S^{n-1}$,

$$\sigma(u \in S^{n-1} : |\langle\theta, u\rangle| \le \frac{c_0}{\sqrt{n}}) = 1/2. \tag{1}$$

We first prove that

Theorem 3. *There exist universal constants c_1, C, c such that for every dimension n, and every $\frac{c_1}{\sqrt{n}} < \varepsilon < 1$, letting $N = Cn\ln n/\varepsilon^2$, if $z_1, \dots, z_N$ are random points on S^{n-1}, then with probability greater than $1 - 2e^{-n\ln(2n+1)}$, the set*

$$\mathcal{K} = \{x \in \mathbb{R}^n : \exists i_1, \dots, i_{N/2} \text{ with } |\langle x, z_{i_j}\rangle| < \frac{c_0}{\sqrt{n}}\}$$

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

Proof. We build two nets, $\mathcal{N}_1$ a $\frac{1}{n}$-net on $(1-\varepsilon)S^{n-1}$ and $\mathcal{N}_2$ a $\frac{1}{n}$-net on $(1+\varepsilon)S^{n-1}$. It is well known that their cardinalities are smaller than $e^{n\ln(2n+1)}$. We define two bodies,

$$\mathcal{K}_1 = \{x \in \mathbb{R}^n : \exists i_1, \dots, i_{N/2} \text{ with } |\langle x, z_{i_j}\rangle| < \frac{c_0}{\sqrt{n}} - \frac{1}{n}\},$$
$$\mathcal{K}_2 = \{x \in \mathbb{R}^n : \exists i_1, \dots, i_{N/2} \text{ with } |\langle x, z_{i_j}\rangle| < \frac{c_0}{\sqrt{n}} + \frac{1}{n}\},$$

and show that with probability greater than $1 - e^{-n\ln n}$, $\mathcal{N}_1 \subset \mathcal{K}_1$ and $\mathcal{N}_2 \cap \mathcal{K}_2 = \emptyset$. This will already imply that $(1-\varepsilon)D_n \subset \mathcal{K} \subset (1+\varepsilon)D_n$, since for $y \in (1-\varepsilon)S^{n-1}$ there will be some $x \in \mathcal{N}_1$ with $|y - x| < \frac{1}{n}$, and so if $|\langle x, z_i\rangle| < \frac{c_0}{\sqrt{n}} - \frac{1}{n}$ for some subset of indices i, then for the same set of indices we also have $|\langle y, z_i\rangle| < \frac{c_0}{\sqrt{n}}$. Similarly, every $y \in (1+\varepsilon)S^{n-1}$ will have an $x \in \mathcal{N}_2$ with $|y - x| < \frac{1}{n}$, and for this x there will be at least $N/2$ indices i for which $|\langle x, z_i\rangle| \ge \frac{c_0}{\sqrt{n}} + \frac{1}{n}$. This implies that for y, for these indices, $|\langle y, z_i\rangle| \ge \frac{c_0}{\sqrt{n}}$, and hence $y \notin \mathcal{K}$. Since $\mathcal{K}$ is star-shaped, this would complete the proof.

Let $x_j \in \mathcal{N}_1$. Then, since $|x_j| = 1 - \varepsilon$,

$$\mathbb{P}[z \in S^{n-1} : |\langle z, x_j\rangle| \le \frac{c_0}{\sqrt{n}} - \frac{1}{n}] = \sigma(u \in S^{n-1} : |\langle\theta, u\rangle| \le \frac{\frac{c_0}{\sqrt{n}} - \frac{1}{n}}{(1-\varepsilon)}) =: p_\varepsilon.$$

Since $\varepsilon > c_1/\sqrt{n}$, and provided we choose $c_1 > 1/c_0$, it is easily seen that $\frac{\frac{c_0}{\sqrt{n}} - \frac{1}{n}}{(1-\varepsilon)} > \frac{c_0}{\sqrt{n}}$, and so $p_\varepsilon > 1/2$. We will later show that $p_\varepsilon - 1/2 \simeq \varepsilon$, but for this stage of the proof it is enough to notice that p_ε is some probability greater than 1/2, and this will give an estimate for $C(\varepsilon)$ depending on p_ε. We use Chernoff's Lemma 2, which says that for at least 1/2 of the N experiments $\{|\langle z_i, x_j\rangle| \le \frac{c_0}{\sqrt{n}} - \frac{1}{n}\}$ to succeed (x_j is fixed and z_i are random, $i = 1, \dots, N$), when the probability of success is p_ε, the chances are high, greater than

$$1 - e^{-NI(p_\varepsilon, 1/2)} = 1 - e^{-N\ln(\frac{1}{4p_\varepsilon(1-p_\varepsilon)})}.$$

Therefore the chances that this will happen simultaneously for all $x_j \in \mathcal{N}_1$ are greater than $1 - e^{n\ln(1+2n)}e^{-N\ln(\frac{1}{4p_\varepsilon(1-p_\varepsilon)})}$. For this to be greater than $1 - e^{-n\ln(2n+1)}$ it suffices that $N > 2n\ln(1+2n)/\ln(\frac{1}{4p_\varepsilon(1-p_\varepsilon)})$.

We now turn to the other side, which is surprisingly similar. Let $x_j \in \mathcal{N}_2$. Then, since $|x_j| = 1+\varepsilon$,

$$\mathbb{P}[z \in S^{n-1} : |\langle z, x_j\rangle| \leq \frac{c_0}{\sqrt{n}} + \frac{1}{n}] = \sigma(u \in S^{n-1} : |\langle \theta, u\rangle| \leq \frac{\frac{c_0}{\sqrt{n}} + \frac{1}{n}}{(1+\varepsilon)}) := q_\varepsilon.$$

We will later show also that $1/2 - q_\varepsilon \simeq \varepsilon$, but for this stage it is enough to notice that it is some probability smaller than 1/2, and this is true provided that $c_1 > 2/c_0$ (similar to the case of $p_\varepsilon - 1/2$). We use again Lemma 2, which says that for at least 1/2 of the N experiments $\{|\langle z_i, x_j\rangle| \leq \frac{c_0}{\sqrt{n}} + \frac{1}{n}\}$ to succeed, when the probability of success is q_ε, the chances are low, smaller than

$$e^{-NI(q_\varepsilon, 1/2)} = e^{-N\ln(\frac{1}{4q_\varepsilon(1-q_\varepsilon)})}.$$

Hence the chances that this will not happen, simultaneously for all $x_j \in \mathcal{N}_2$, are greater than $1 - e^{n\ln(1+2n)}e^{-N\ln(\frac{1}{4q_\varepsilon(1-q_\varepsilon)})}$. We will thus assume also that $N > 2n\ln(1+2n)/\ln(\frac{1}{4q_\varepsilon(1-q_\varepsilon)})$.

All that is left is to estimate $(p_\varepsilon - 1/2)$ and $(1/2 - q_\varepsilon)$. The quantity $(p_\varepsilon - 1/2)$ is twice the volume, on the sphere, between two parallel hyperplanes, one at distance $\frac{c_0}{\sqrt{n}}$ from the origin and one at distance $(\frac{c_0}{\sqrt{n}} - \frac{1}{n})/(1-\varepsilon)$. The quantity $(1/2 - q_\varepsilon)$ is twice the volume, on the sphere, between two parallel hyperplanes, one at distance $\frac{c_0}{\sqrt{n}}$ from the origin and one at distance $(\frac{c_0}{\sqrt{n}} + \frac{1}{n})/(1+\varepsilon)$. To estimate these volumes from below, we use the fact that $\varepsilon > c_1/\sqrt{n}$ to obtain that the distance between the two hyperplanes is bounded from below in both cases by $(c_1c_0/2 - 1)/n$. We now use Fubini's theorem and the volume of a section of the sphere to see that $(p_\varepsilon - 1/2)$ and $(1/2 - q_\varepsilon)$ are bounded from below by $c_2\varepsilon$ for some universal c_2 (which depends linearly on our choice of $c_1 > 2/c_0$). Using this to analyze the estimates for N, we arrive at the desired result. □

Proof of Theorem 1. As before, we define

$$\mathcal{K} = \{x \in \mathbb{R}^n : \exists i_1, \ldots, i_{N/2} \text{ with } |\langle x, z_{i_j}\rangle| < \frac{c_0}{\sqrt{n}}\},$$

however this time $N = C(\varepsilon)n$, with $C(\varepsilon) = C'\ln(1/\varepsilon)/\varepsilon^2$, where C' is a universal constant.

We first show that with high probability $\mathcal{K} \subset (1+\varepsilon)D_n$. We abbreviate $C(\varepsilon) = C$. Cover $(1+\varepsilon)S^{n-1}$ by balls of radius δ to be determined later. As is well known, such a covering exists with cardinality less than $(1+\frac{2(1+\varepsilon)}{\delta})^n$. Consider one of these balls, $B(x_0, \delta)$, with $|x_0| = (1+\varepsilon)$. Let $c_1(\eta)$ be the constant such that

$$\mathbb{P}[z \in S^{n-1} : |\langle z, \frac{x_0}{|x_0|}\rangle| < c_1(\eta)\frac{1}{\sqrt{n}}] = 1/2 - \eta.$$

This constant will depend on n. However, similar to c_0 in (1), asymptotically as $n \to \infty$, it depends only on η. Also, clearly, as η tends to 0, the constant $c_1(\eta)$ tends to c_0. After tossing N points z_i in S^{n-1}, there is a high probability that more than $(1/2 + \eta/2)N$ of them will satisfy $\{|\langle z, x_0\rangle| \geq (1+\varepsilon)c_1\frac{1}{\sqrt{n}}\}$. More precisely,

$$\begin{aligned}\mathbb{P}[z_1, \ldots, z_N \in S^{n-1} : &\text{for } (1/2+\eta/2)N \text{ of the} \\ &\text{indices } i,\ |\langle z_i, x_0\rangle| \geq (1+\varepsilon)c_1(\eta)\frac{1}{\sqrt{n}}] \quad \geq \quad 1 - c(\eta)^N,\end{aligned}$$

where $c(\eta) = e^{-[(1/2-\eta/2)\ln(\frac{1/2-\eta/2}{1/2-\eta}) + (1/2+\eta/2)\ln(\frac{1/2+\eta/2}{1/2+\eta})]}$.

If we think of the random z_i as the rows of a matrix, a concentration result for the norm of such a matrix applies. We know that (see [MP], Lemma 1 (ii): the same proof applies for this version)

$$\mathbb{P}[\forall y, |y| \geq \frac{1}{3(\sqrt{C}+1)}(\sum_{i=1}^{N}\langle z_i, y\rangle^2)^{1/2}] \geq 1 - e^{-N}. \tag{2}$$

If $x \in \mathcal{K}$, then for half of the indices i one has $|\langle x, z_i\rangle| < \frac{c_0}{\sqrt{n}}$. In particular, there are $N\eta/2$ indices i for which, at the same time, we also have $|\langle z_i, x_0\rangle| > (1+\varepsilon)c_1(\eta)\frac{1}{\sqrt{n}}$, and thus

$$\langle z_i, x_0 - x\rangle^2 \geq ((1+\varepsilon)c_1(\eta) - c_0)^2/n.$$

Joining these two facts together, one can use (2) to see that for a specific x_0, with probability $1 - c(\eta)^N - e^{-N}$, for all $x \in \mathcal{K}$

$$|x_0 - x| \geq \frac{1}{3(\sqrt{C}+1)}\sqrt{(\eta/2)N((1+\varepsilon)c_1 - c_0)^2/n} \geq \frac{\sqrt{\eta}}{9}((1+\varepsilon)c_1(\eta) - c_0).$$

Since $c_1(\eta) \to c_0$ as $\eta \to 0$, for each ε there is an $\eta = \eta(\varepsilon)$ such that $(1+\varepsilon)c_1(\eta) - c_0 = \theta(\varepsilon) > 0$. The above implies that for this η

$$\mathcal{K} \cap B(x_0, \frac{\sqrt{\eta}}{9}\theta(\varepsilon)) = \emptyset.$$

If, to begin with, we choose $\delta = \frac{\sqrt{\eta}}{9}\theta(\varepsilon)$, then having this for all of the balls in the covering will imply that $\mathcal{K}$ does not intersect $(1+\varepsilon)S^{n-1}$, and thus $\mathcal{K} \subset (1+\varepsilon)D_n$.

We have to calculate the probability of this event and show that it is large (and in particular, positive) for our choice of $C = C(\varepsilon)$. The probability of the complement event is at most

$$\left(1 + \frac{2(1+\varepsilon)}{\frac{\sqrt{\eta}}{9}\theta(\varepsilon)}\right)^n (c(\eta)^{Cn} + e^{-Cn}). \tag{3}$$

Since all the constants (η, $c(\eta)$, $\theta(\eta)$) depend only on ε, it is clear that if C is big enough (depending only on ε), this probability is very small. What is left is to compute the dependency of $C = C(\varepsilon)$ on ε. We begin by considering $\eta(\varepsilon)$, which has to satisfy that $(1+\varepsilon)c_1(\eta) - c_0 = \theta(\varepsilon) > 0$. The same reasoning as in the last paragraph of the proof of Theorem 3 shows that we need to take $\eta = c_2\varepsilon$ (where c_2 is a universal constant; remember that so is c_0) and then get that $\theta(\varepsilon) \geq c_3\varepsilon$. Taking the second order in η we see that $c(\eta) = e^{-\eta^2/2+O(\eta^3)}$. Therefore, for $\eta = c_2\varepsilon$, we have $c(\eta) \leq e^{-c_4\varepsilon^2}$. The probability in (3) can thus be bounded by

$$\left(\frac{c_5}{\varepsilon^{3/2}}\right)^n (e^{-Cc_4\varepsilon^2 n} + e^{-Cn})$$

for universal c_4, c_5. We can see that the dependency of C on n and ε which we get with this method is $C = C'\ln(1/\varepsilon)/\varepsilon^2$ for a universal C'.

We now have to verify the other side of the statement in the theorem, which is that $(1-\varepsilon)D_n \subset \mathcal{K}$. We again abbreviate $C(\varepsilon) = C$. This time cover $(1-\varepsilon)S^{n-1}$ by $(1 + \frac{2(1-\varepsilon)}{\delta})^n$ balls of radius δ. Consider one of these balls, $B(x_0, \delta)$, with $|x_0| = (1-\varepsilon)$. Let $c_1(\eta)$ this time denote the constant (depending asymptotically only on η, and tending to c_0 as $\eta \to 0$) such that

$$\mathbb{P}[z \in S^{n-1} : |\langle z, \frac{x_0}{|x_0|}\rangle| < c_1(\eta)\frac{1}{\sqrt{n}}] = 1/2 + \eta.$$

Then, after tossing $N = Cn$ points z_i in S^{n-1}, there is a high probability that no more than $1/2 - \eta/2$ of them will satisfy $|\langle z, x_0\rangle| \geq (1-\varepsilon)c_1\frac{1}{\sqrt{n}}$. More precisely,

$$\mathbb{P}[z_1, \ldots, z_N \in S^{n-1} : \text{for } 1/2 + \eta/2 \text{ of the indices } i,\ |\langle z_i, x_0\rangle| \leq (1-\varepsilon)c_1(\eta)\frac{1}{\sqrt{n}}] \quad \geq \quad 1 - c(\eta)^N,$$

where as before $c(\eta) = e^{-[(1/2-\eta/2)\ln(\frac{1/2-\eta/2}{1/2-\eta}) + (1/2+\eta/2)\ln(\frac{1/2+\eta/2}{1/2+\eta})]}$.

We again use that $\mathbb{P}[\forall y, (\sum_{i=1}^N \langle z_i, y\rangle^2)^{1/2} \leq 3(\sqrt{C}+1)|y|] \geq 1 - e^{-N}$. If $x \notin \mathcal{K}$, then for half of the indices i one has $|\langle x, z_i\rangle| > \frac{c_0}{\sqrt{n}}$. In particular, there are $N\eta/2$ indices i for which, at the same time, we also have $|\langle z_i, x_0\rangle| < (1-\varepsilon)c_1(\eta)\frac{1}{\sqrt{n}}$, and thus

$$\langle z_i, x - x_0\rangle^2 \geq (c_0 - (1-\varepsilon)c_1(\eta))^2/n.$$

Thus, for a specific x_0, with probability $1 - c(\eta)^N - e^{-N}$, for all $x \notin \mathcal{K}$

$$\sqrt{(\eta/2)N(c_0 - (1-\varepsilon)c_1)^2/n} \leq 3(\sqrt{C}+1)|x - x_0|.$$

In particular, $x \notin B(x_0, (\sqrt{\eta}/9)(c_0 - (1-\varepsilon)c_1))$. Since $c_1(\eta) \to c_0$ as $\eta \to 0$, for each ε there is an $\eta = \eta(\varepsilon)$ such that $c_0 - (1-\varepsilon)c_1(\eta) = \theta(\varepsilon) > 0$. The above implies that for this η

$$\mathcal{K} \supset B(x_0, \frac{\sqrt{\eta}}{9}\theta(\varepsilon)).$$

We choose $\delta = \frac{\sqrt{\eta}}{9}\theta(\varepsilon)$ and then having this for all of the balls in the covering implies that $\mathcal{K}$ includes the whole of $(1-\varepsilon)D_n$.

We should calculate the probability of this event and show that it is large, and then check the dependency of C on ε. Both things are done in exactly the same way as in the proof of the first part, and we omit the details. □

Some other results which can be obtained by this method

The same method of proof can be used in many other applications. We list below several other theorems which can be proven using Chernoff bounds in a geometric context. Most of them share the same basic idea: If one looks for a lower bound for $\frac{1}{N}S_N = \frac{1}{N}\sum_{i=1}^N X_i$, one can define the Bernoulli random variables Z_i to equal 1 if $X_i \geq \varepsilon$ and 0 otherwise. If X_i are positive random variables, then the probability that $\frac{1}{N}S_N \geq \beta\varepsilon$ is at least the probability that βN of the Z_i's are 1. When $\beta < \mathbb{P}[X_i \geq \varepsilon]$, this probability can be estimated using Chernoff's Lemma 2. We remark that upper bounds for such sums are usually much easier.

The first theorem we wish to mention is in the spirit of [S1], and shows how little randomness is needed to realize an n-dimensional euclidean section of $\ell_1^{(1+\delta)n}$ by a sign-matrix.

Theorem 4. *There exist universal constants c_1, c' and C such that the following holds. Let $n = 2^m$, $\delta > 0$, and denote $N = (1+\delta)n$ and $c(\delta) = c_1\delta/(\ln(1/\delta))$. If we denote by A the $n(1+\delta) \times n$ matrix whose first n rows are a Walsh sign matrix and the last δn rows are random sign-vectors $\varepsilon(j) \in \{-1,1\}^n$, then, with probability $1 - e^{-c'\delta n}$, we will have for every $x \in \mathbb{R}^n$*

$$c(\delta)|x| \leq \|Ax\|_{L_1} \leq (1 + \sqrt{\delta}C)|x|, \tag{4}$$

(where $|x| = (\sum_{i=1}^n x_i^2)^{1/2}$ and $||y||_{L_1} = \frac{1}{N}\sum_{i=1}^N |y_i|$).

These types of questions can be viewed in two ways. One is as realizing a euclidean section of ℓ_1 with a sign matrix and the second is as getting a form of Khinchine's inequality which does not involve *all* 2^n sign-vectors but just $(1+\delta)n$ of them. The first result in this direction was by Schechtman [S2], where he showed that for δ bigger than some universal constant this probability is exponentially large. Then, together with Johnson in [JS] they showed that for any $\delta > 0$ there exists a sign-matrix with this property, which means that the probability is positive. The question remained whether this probability, for any $\delta > 0$, is large, i.e., does a random sign matrix satisfy inequality (4) with, perhaps, a different dependency of the constants on δ. In a recent theorem of Litvak, Pajor, Rudelson, Tomczak-Jaegermann and Vershynin [LPRTV], they showed (following developments described in [LPRT]) a general result for bodies with finite volume ratios, a particular case of which is that there exists a constant $0 < c(\delta)$, such that for $N = (1+\delta)n$ random sign vectors $\varepsilon(1), \dots, \varepsilon(N) \in \{-1,1\}^n$, with probability greater than $1 - e^{-c'n}$, one has for every $x \in \mathbb{R}^n$

$$c(\delta)|x| \le \frac{1}{N}\sum_{j=1}^{N} |\langle \varepsilon(j), x\rangle| \le C|x|. \tag{5}$$

Both C and c' are universal constants. (In their general result the middle ℓ_1 norm was replaced by an arbitrary norm with bounded volume ratio, and the constants can depend on the volume ratio.)

With the method using Chernoff-type estimates we are also able to prove (5). Our proof provides a constant $c(\delta) \ge (c_1\delta)^{1+1/\delta}$, where c_1 is an absolute constant. The proof from [LPRTV] (while optimal in the *general* context) gives only a slightly better behavior of the constant: $c(\delta) \ge c_1^{1+1/\delta}$. This exponentially bad dependence should be compared with the good, polynomial, behavior of the constant in Theorem 4.

The direction of constructing good euclidean sections of ℓ_1, which are random in some sense, can be continued in a different way. With the method we described we can prove the following theorem, in which the condition of independence of the different random matrix coefficients is relaxed. We use the notion of an "isotropic body": A body K is called isotropic if it satisfies $Vol(K) = 1$, $\int_K x = 0$ and, most importantly, for every $\theta \in S^{n-1}$ the integral $\int_K \langle x, \theta\rangle^2$ is a constant independent of θ, depending only on K, which is called the (square of the) isotropic constant of K and denoted L_K^2. (It is easy to check that every body has a linear image which is isotropic. In other words, saying that the body is in isotropic position only means that we identify the right euclidean structure.)

We can prove the following theorem, in which the random matrix of signs is replaced by a random matrix whose rows are random points in an isotropic convex body.

Theorem 5. *For any $\delta > 0$ there exist constants $0 < c(\delta)$, depending only on δ and universal constants $0 < c', C < \infty$ such that for any convex body K in isotropic position, with probability greater than $1 - e^{-c'n}$ we have that*

$$c(\delta)L_K|x| \le \frac{1}{N}\sum_{j=1}^{N} |\langle z_j, x\rangle| \le CL_K|x|,$$

where $N = (1+\delta)n$, and where z_j are chosen independently and uniformly inside the body K.

Again, the constant $c(\delta) = (c_1\delta)^{1+1/\delta}$, for a universal c_1.

The motivation and the details of proofs for these theorems, and other applications of the method, will appear in [AFM].

References

[AFM] S. Artstein-Avidan, O. Friedland, V. Milman, *Geometric Applications of Chernoff-type Estimates*, to appear in Springer Lecture Notes, GAFA Seminar, 2004–2005.

[B] A. R. Barron, *Universal approximation bounds for superpositions of a sigmoidal function*, IEEE Trans. Inform. Theory 39 (1993), 930–945. MR1237720 (94h:92001)

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

[C] G. Cybenko, *Approximation by superpositions of sigmoidal functions*, Proc. of the 1994 IEEE-IMS Workshop on Info. Theory and Stat. (1994).

[HR] T. Hagerup and C. Rüb, *A guided tour of Chernoff bounds*, Inform. Process. Lett. 33 (1990) no. 6, 305–308. MR1045520 (91h:60022)

[HSW] K. Hornik, M.B. Stinchcombe and H. White, *Multi-layer feedforward networks are universal approximators*, Neural Networks 2 (1988), 359–366.

[JS] W.B. Johnson and G. Schechtman, *Very tight embeddings of subspaces of L_p, $1 \le p < 2$, into ℓ_p^n*, Geom. and Funct. Anal. (GAFA), 13, no. 4 (2003), 845–851. MR2006559 (2004h:46014)

[LPRT] A.E. Litvak, A. Pajor, M. Rudelson, and N. Tomczak-Jaegermann, *Smallest singular value of random matrices and geometry of random polytopes*, preprint.

[LPRTV] A.E. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, and R. Vershynin, *Random Euclidean embeddings in spaces of bounded volume ratio*, C. R. Math. Acad. Sci. Paris 339 (2004), no. 1, 33–38. MR2075229

[MP] V.D. Milman and A. Pajor, *Regularization of star bodies by random hyperplane cut off*, Studia Methematica 159 (2) (2003), 247–261. MR2052221 (2005e:52009)

[S1] G. Schechtman, *Special orthogonal splittings of L_1^{2k}*, Israel J. Math. 139 (2004), 337–347. MR2041797 (2005a:46036)

[S2] G. Schechtman, *Random embeddings of euclidean spaces in sequence spaces*, Israel Journal of Mathematics 40, no. 2 (1981), 187–192. MR0634905 (83a:46032)

Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, New Jersey 08544-1000 – and – School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton, New Jersey 08540

E-mail address: artstein@princeton.edu

School of Mathematical Science, Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel

E-mail address: omerfrie@post.tau.ac.il

School of Mathematical Science, Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel

E-mail address: milman@post.tau.ac.il