Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Functional Analysis 235 (2006) 297–329
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

Logarithmic reduction of the level of randomness in some probabilistic geometric constructions ☆

S. Artstein-Avidan [a,b,*], V.D. Milman [c]

[a] *Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544-1000, USA*
[b] *School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA*
[c] *School of Mathematical Science, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel*

Received 6 November 2005; accepted 6 November 2005

Available online 19 January 2006

Communicated by J. Bourgain

Abstract

Many of the surprising phenomena occurring in high dimensions are proved by use of probabilistic arguments, which show the existence of organized and regular structures but do not hint as to where exactly do these structures lie. It is an intriguing question whether some of them could be realized explicitly. In this paper we show that the amount of randomness used can be reduced significantly in many of these questions from asymptotic convex geometry, and most of the random steps can be substituted by completely explicit algorithmic steps. The main tool we use is random walks on expander graphs.

Keywords: Randomness reduction; Explicit constructions; Sections of ℓ_1; Asymptotic geometric analysis

☆ This research was partially supported by BSF grant 2002-006, the first named author was also supported by the National Science Foundation under agreement No. DMS-0111298. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.
* Corresponding author.
E-mail addresses: artstein@princeton.edu (S. Artstein-Avidan), milman@post.tau.ac.il (V.D. Milman).

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2005.11.003

A well-known phenomenon in high-dimensional geometry and in many other fields of mathematics and computer science is the existence of structures which are obtained probabilistically, but for which an explicit construction is unknown. In most cases the situation is as follows: with the natural probability measure on the set of possible constructions, the probability that a random one satisfies some desired properties is very high, close to one, but at the same time we cannot point at even one specific (deterministic) construction with these properties.

In this paper we will deal with geometric situations, which arise in the framework of asymptotic geometric analysis. Thus, usually, we have a linear normed space whose dimension is assumed to be large, and any constants appearing, unless specifically stated, do not depend on the dimension.

In addition to the obvious interest in finding a concrete structure with some desired properties, the existence of which is guaranteed from probabilistic considerations, we expect that knowing exactly where "good" structures are hiding will reveal some geometric properties of the space in question.

One such example is a section of the cross-polytope (the unit ball of ℓ_1^n) which is of proportional dimension and isomorphic to a Euclidean ball up to a constant independent of the dimension n. It has been known from the early works such as [12,18] that a *random* section, with high probability, satisfies this property (random here with respect to the Haar measure on the Grassmanian of all, say, $(n/2)$-dimensional subspaces of $\mathbb{R}^n$). However, there is no known explicit construction of such a section of ℓ_1^n. We return to this example in Section 6.

One important case in which an explicit construction *is* known is that of a subspace of a projection of the simplex, given in [8]. They use in an essential way the non-symmetric structure of the simplex, and give a deterministic algorithm which finds a concrete section of a concrete projection of the simplex in dimension which is a fifth of the original dimension, and which is close to Euclidean, with an exceptionally good isomorphism constant. Thus, one more goal which one might expect to achieve based on this one result, is that in explicit or near explicit constructions the isomorphism constants which one gets might be much better than what is known about random constructions.

In this note we will present several examples in which although one does not find an explicit construction, one is able to reduce significantly the "amount" of randomness used to obtain a certain structure with desired properties. The examples we discuss in this note include:

- Global Dvoretzky-type theorem;
- ZigZag bodies;
- Sections of ℓ_1^n (and of other finite volume ratio bodies);
- The low M^* estimate;
- The Quotient of a Subspace theorem.

In the first example we reduce randomness by means of an iteration procedure. In the rest of the examples we use a derandomization method relying on random walks in expander graphs. Thus, actually, our understanding of explicitness is an explicit algorithm which leads to the desired result with a small number of steps. We note that this is the same

understanding of algorithmic explicitness as in the paper [8] mentioned above, where the authors construct an algorithm to derive a subspace of a quotient of the simplex. However, their algorithm (which is completely explicit) is based on linear programming. Our algorithms are based on expander graphs, and need some randomness, though much less than the standard techniques require.

To do this we use tools developed in computer science to investigate the pseudo-random properties of walks on expander graphs. However, usually these tools are not enough and we have to develop them further. In particular, Chernoff bounds for this type of walks (to be later described) were well known and used by many authors, whereas Bernstein type theorems such as Theorems 12 and 14 were to our knowledge previously unknown, and play an essential role in our proofs.

We also had to, in some cases, adapt the geometric statements and proofs to the kind which allows derandomization. In particular, provide some variations on the standard theorems mentioned above where the spaces are either spanned by or are the kernels of families of sign-vectors.

Notation. We use S^{n-1} to denote the Euclidean sphere or radius 1 in $\mathbb{R}^n$ and D_n to denote the Euclidean ball. The Euclidean norm of a vector x we denote simply by $|x|$, so $D_n = \{x \in \mathbb{R}^n\colon\ |x| \leqslant 1\}$. For a symmetric convex body K we denote $M = M(K) = \int_{S^{n-1}} \|u\|_K \, d\sigma(u)$, where $\|\cdot\|_K$ is the norm whose unit ball is K. We denote $M^*(K) = M(K^\circ)$ where K° is the dual body for K, that is $K^\circ = \{u\colon\ \sup_{x\in K}\langle u, x\rangle \leqslant 1\}$. One may also write $M^*(K) = \int_{S^{n-1}} \sup_{x\in K}\langle u, x\rangle \, d\sigma(u)$, and it is called half the mean width of K. We usually use d for the half-diameter (or radius) of K, and b for the half-diameter of its dual, that is, $\frac{1}{b} D \subseteq K \subseteq d\,D$.

We denote by $O(n)$ the group of orthogonal rotations in $\mathbb{R}^n$.

1. Global Dvoretzky

Given a convex body $K \subset \mathbb{R}^n$, the global Dvoretzky-type theorem, which first appeared in [9], in a non-explicit form, and later explicitly in [30], states that the Minkowski average of $C(d/M^*)^2$ random rotations of the body K is isomorphic to a Euclidean ball of radius M^*. More precisely

Theorem 1. *There exist universal constants c, c', C and C' such that for every symmetric convex body $K \subset \mathbb{R}^n$ satisfying $K \subseteq d\,D$, and with $M^* = M^*(K)$, with probability $1 - e^{-c'n}$, the $N = C'(d/M^*)^2$ random orthogonal transformations $U_1, \ldots, U_N \in O(n)$ satisfy for every $x \in \mathbb{R}^n$ that*

$$cM^* D_n \subset \frac{1}{N}\sum_{i=1}^{N} U_i K \subset CM^* D_n.$$

Restated in dual form, this theorem reads as follows.

Theorem 1′. *There exist universal constants c, c', C and C' 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 $1 - e^{-c'n}$, the $N = C'(b/M)^2$ random orthogonal transformations $U_1, \ldots, U_N \in O(n)$ satisfy for every $x \in \mathbb{R}^n$ that*

$$cM|x| \leqslant \frac{1}{N} \sum_{i=1}^{N} \|U_i x\|_K \leqslant CM|x|.$$

The "level of randomness" in this result is $N = C'(b/M)^2$ random orthogonal transformations. This bound for the number of rotations involved is optimal up to the constants, see [30]. Moreover, in [30] it is shown that this bound on the number of rotations is optimal regardless of the way they are constructed, that is, also in the non-random setting, one cannot do with less. In this section we explain how to reduce the level of randomness in this result to $N = C_1 + C_2 \log(b/M)^2$. The number of operators will of course be larger, but they will be explicitly constructed using just N such *random* operators.

To formulate the result it is notationally convenient to denote $U_i^0 = Id$ the identity operator, and $U_i^1 = U_i$. We prove

Theorem 2. *There exist universal constants c, c', C, C_1 and C_2 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 $1 - e^{-c'n}$, the $N = C_1 + C_2 \log(b/M)^2$ random orthogonal transformations $U_1, \ldots, U_N \in O(n)$ satisfy for every $x \in \mathbb{R}^n$*

$$cM|x| \leqslant \frac{1}{N} \sum_{\delta \in \{0,1\}^N} \left\| \prod_{i=1}^{N} U_i^{\delta_i} x \right\|_K \leqslant CM|x|.$$

In fact, this is true with C_2 as close to 1 as we want, but then C_1 grows and c' decreases.

Another way to state this result is as follows:

In the same notation as above, with probability $1 - e^{-c'n}$, defining

$$\|x\|_1 = \|x\|_K \quad \text{and} \quad \|x\|_j = \frac{\|x\|_{j-1} + \|U_j x\|_{j-1}}{2}$$

one has for every $x \in \mathbb{R}^n$ that

$$cM|x| \leqslant \|x\|_N \leqslant CM|x|.$$

We will actually prove it in its dual form which reads

Theorem 2′. *There exist universal constants c, c', C, C_1 and C_2 such that for every symmetric convex body $K \subset \mathbb{R}^n$ satisfying $K \subseteq dD$, and with $M^* = M^*(K)$, with prob-*

ability $1 - e^{-c'n}$, *the* $N = C_1 + C_2 \log(d/M^*)^2$ *random orthogonal transformations* $U_1, \ldots, U_N \in O(n)$ *satisfy for every* $x \in \mathbb{R}^n$ *that, denoting* $K_0 = K$ *and*

$$K_j = \frac{U_j K_{j-1} + K_{j-1}}{2},$$

we have

$$cM^* D_n \subset K_N \subset CM^* D_n.$$

Again, this is true with C_2 as close to 1 as we want, but then C_1 grows and c' decreases.

The idea is that averaging two copies of the body K, reduces the diameter of the resulting body by approximately $\sqrt{2}$, provided the parameter $(d/M^*)^2$ is large. The mean-width M^* of the new body remains unchanged. Repeating this t times, we get that the diameter d_t of the new body is approximately $d/2^{t/2}$ (as long as $(d_t/M^*)^2$ is large). After at most a logarithmic number of steps, we arrive at a body with bounded d_t/M^*, and then using Theorem 1 we have that the average of a bounded number of random rotations of this new body is enough to reach an isomorphic Euclidean ball.

So we see that first we need the following lemma, about the decrease in diameter after one step of this procedure.

Lemma 3. *For any* $0 < \varepsilon < 1$ *there exists constants* $C(\varepsilon), c_\varepsilon$ *such that for a symmetric convex body* K, *if* $C(\varepsilon) < \operatorname{diam}(K)/M^*(K)$ *then for a random* $U \in O(n)$ *we have with probability greater than* $1 - e^{-c_\varepsilon n}$ *that*

$$\operatorname{diam}\left(\frac{K + UK}{2}\right) \leqslant \frac{(1+\varepsilon)}{\sqrt{2}} \operatorname{diam}(K).$$

One may take, for example, $C(\varepsilon) = 40/\varepsilon^2$ *and* $c_\varepsilon = \varepsilon^2/16$.

Proof. First use Sudakov's inequality to cover K with T Euclidean balls of radius αd, for $\alpha < 1$ to be chosen later, where $T \leqslant e^{2n(M^*/(\alpha d))^2}$ (the constant 2 can actually be replaced by a constant tending to 1 as $T \to \infty$). Denote the centers of this covering by x_i, $i = 1, \ldots, T$. We will estimate the probability on $U \in O(n)$ for which $|\langle x_i, Ux_j\rangle| \leqslant \delta d^2$ for every $i = 1, \ldots, k$ and $j = 1, \ldots, k$, where δ will be chosen later. Indeed, for a fixed i, j one has

$$\mathbb{P}\big[U\colon \big|\langle x_i, Ux_j\rangle\big| \leqslant \delta d^2\big] \geqslant \sigma\left(y \in S^{n-1}\colon \left|\left\langle \frac{x_i}{|x_i|}, y\right\rangle\right| \leqslant \delta\right) \geqslant 1 - e^{-\delta^2 n/2}.$$

We have at most $e^{4n(M^*/(\alpha d))^2}$ such pairs, so, as long as, say, $e^{\delta^2 n/2} \geqslant e^{8n(M^*/(\alpha d))^2}$, the probability of the operators U for which this is true for all pairs (i, j) is greater than $1 - e^{-\delta^2 n/4}$. This condition amounts to $\delta \geqslant 4(M^*/(\alpha d))$.

Let $z \in (K + UK)/2$, that is, $z = (y_1 + Uy_2)/2$ where $y_1, y_2 \in K$. Each of the points y_1, y_2 has a closest point in the net: $y_1 = x_{i_1} + \alpha d\theta_1$ and $y_2 = x_{i_2} + \alpha d\theta_2$, where $\theta_1, \theta_2 \in D_n$. We now have

$$\begin{aligned}\left|\frac{y_1+Uy_2}{2}\right|^2 &= \left|\frac{x_{i_1}+\alpha d\theta_1+Ux_{i_2}+\alpha dU\theta_2}{2}\right|^2\\ &= \frac{|x_{i_1}|^2}{4}+\frac{|x_{i_2}|^2}{4}+\frac{\langle x_{i_1},Ux_{i_2}\rangle}{2}+\frac{|\alpha d(\theta_1+U\theta_2)|^2}{4}\\ &\quad+\frac{\langle x_{i_1}+Ux_{i_2},\alpha d(\theta_1+U\theta_2)\rangle}{2}\\ &\leqslant d^2\left(\frac{1}{2}+\frac{\delta}{2}+\alpha^2+2\alpha\right).\end{aligned}$$

We choose, for example, $\alpha=\varepsilon/5$, $\delta=\varepsilon/2$, so that we get the inequality

$$\left|\frac{y_1+Uy_2}{2}\right|\leqslant d\sqrt{\frac{1}{2}+\varepsilon}\leqslant d\frac{1+\varepsilon}{\sqrt{2}},$$

and the condition above amounts to $\varepsilon^2/10\geqslant 4(M^*/d)$, so that $C(\varepsilon)=40/\varepsilon^2$ and $c_\varepsilon=\varepsilon^2/16$ works. Notice that in fact we can do better, we can choose say $\alpha=\delta=2\sqrt{M^*/d}$ and get an even more isometric result (with a worse probability estimate) in the domain where M^*/d is small. □

Proof of Theorem 2′. Denote by d_j the diameter of the body K_j. Notice that $M^*(K_j)=M^*(K)$. The lemma promises us that as long as $d_j/M^*>C(\varepsilon)$ we have with high probability that

$$d_{j+1}\leqslant\frac{1+\varepsilon}{\sqrt{2}}d_j.$$

Thus, as long as $d_j/M^*>C(\varepsilon)$ we have that

$$d_j\leqslant\left(\frac{1+\varepsilon}{\sqrt{2}}\right)^j d.$$

We fix, say, $\varepsilon=0.1$ and take $j_0=C_2\log(d/M^*)$, for $C_2=\ln(\sqrt{2}/1.1)$. There are two options,

$$\text{either}\quad\frac{d_{j_0}}{M^*}\leqslant C(\varepsilon),\quad\text{or}\quad d_{j_0}\leqslant\left(\frac{1.1}{\sqrt{2}}\right)^{C\log(d/M^*)}d=M^*.$$

In fact, this second possibility is not realistic since this would mean K_j is an exact ball. Therefore $(d_j/M^*)\leqslant C(\varepsilon)$ and we may use Theorem 1 to deduce that with high probability C_1 random rotations we transform K_{j_0} into an isomorphic Euclidean ball. The interplay between the choice of ε (which determines C_2) and the resulting $C(\varepsilon)$ (which in turn determines C_1) is clear from the proof. Estimating the probability is straightforward, we have $1-[C_2\log(d/M^*)]e^{-c_\varepsilon n}-e^{-c'n}$ which is less than $1-e^{-c''n}$ for a suitable c'' (where we use the fact that d/M^* is always less than $\sqrt{n}$, for example).

Note that we proved in fact that after some $N_1 = C_2 \log(d/M^*)$ steps we may take the usual sum of C_1 more random rotations of the attained body and get a body isomorphic to Euclidean. This is slightly different from the way we stated the theorem, to unify notation. However, as stated in the theorem, with the more complicated sum structure, the isomorphism constants can be only better. □

2. Some background on expanders

In the rest of this paper, we use for derandomization a technique that we borrow from computer science, namely the use of expander graphs to generate pseudo-randomness. This is done by exchanging random choices of vectors with taking a random walk along the edges of a graph whose vertices are indexed by the possible vectors. If we put a complete graph on the set of possible vectors, then taking a random walk on this graph, which starts at a uniformly random vertex, goes at each step, randomly, to one of its neighbors, and selecting as the final set of vectors all the vectors the walk has stepped on, is of course the same as picking the vectors uniformly and independently at random.

However, when taking a random walk on a graph with lower degree, the vectors you get are not i.i.d., and there are strong correlations among them. The gain is the amount of randomness used, which is much smaller: one needs to pick the starting vector at random, but for each subsequent step just $\log d$ bits are needed to chose between its d neighbors. This is true at least if these neighbors are easily "labeled," so that with $\log d$ random bits you can choose one of the neighbors at random.

The pseudo-randomness one gets, namely the properties of this correlated sequence which are similar to properties of i.i.d. random vectors, depend on the properties of the graph chosen and imposed (somewhat artificially) on the points. The best choice seems to be an *expander* of degree d, as is demonstrated, for example, in Theorem 7 below. An *expander* graph is a graph whose normalized adjacency matrix has a large spectral gap, that is, a relatively small second eigenvalue. For background and related material see, e.g., [3,4,20,33]. We remark that explicit constructions for expanders are well known in the theory. Perhaps the example which most easy to visualize is the Cayley graph of the group $SL_3(p)$ with two generators. The vertices of the graph are indexed by the elements of $SL_3(p)$ (there are $(p^3-1)\cdot(p^2-1)(p-1)p^2$ of them) and each vertex has three neighbors corresponding to the generators and their inverses. The fact that this graph is a good expander follows from the Kazhdan property (T), for details see [3,22]. The idea of using a random walk on an expander graph to simulate true uncorrelated randomness goes back to [1], see also [11,15].

We usually use as our space of possible vectors the set $\{-1,1\}^n$, that is, a set with 2^n points. The expander, if taken to be $SL_3(p)$ for some appropriate $p \simeq 2^{n/8}$, may have a slightly larger number of points, but we may still index 2^n of them by the vertices of the cube. In the procedure of a random walk to be described precisely later on in the note, we ignore all steps which happen to be in the remaining unindexed set. The unindexed set can be assumed to have exponentially small probability. This follows from a theorem of Hoheisel from 1930 about primes in "short" intervals, namely that there exists $a < 1$ such that for sufficiently large x there is a prime in the interval $[x; x + x^a]$. (We thank

Z. Rudnick for this reference.) See, e.g., [17, Chapter 10.5]. Thus the unindexed set will not change any of the estimates below, and we ignore this point hereafter, assuming the expander has exactly 2^n points.

Also, we sometimes use in what follows expanders of high degree, which can be constructed either by taking the power of a given expander of low degree (that is, the new edges will be walks of some length in the first expander) or by other explicit constructions which are known. The first explicit construction of an expander is that of Margulis [22], which is followed by [3]. For more details and examples see [4,21,23] and the books specifically on this subject [20,33].

3. ZigZag bodies

In a recent paper of the authors and O. Friedland [7] we addressed the question of approximating the Euclidean ball D_n in $\mathbb{R}^n$ by a simpler set, which we called a ZigZag body, and is described as follows. Given a set of N inequalities, and a number $k \leqslant N$, the associated ZigZag body consists of all points satisfying no less than k of the N inequalities. (Notice that it is usually not a convex set.) We learned of this approximation from a paper by Barron and Cheang [10], 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). Using a new approach involving Chernoff's bounds we improved, in the case of high dimension, their estimate to $N = Cn\ln(1/\varepsilon)/\varepsilon^2$ linear inequalities (again, for a universal C), and we use $k = N/2$. The formulation of our result is given in the following theorem (see [7] for a detailed proof).

Theorem 4. *There exists 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*

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

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 $\sigma(u \in S^{n-1} \colon |\langle \theta, u\rangle| \leqslant c_0/\sqrt{n}) = 1/2$ for some $\theta \in S^{n-1}$.

We would like to find a procedure to decrease the level of randomness in this theorem as well, and build a ZigZag set which is close to the Euclidean ball, using as few random bits as possible. The procedure consists of two steps. In the first one, we prove a similar theorem where instead of random sphere vectors we use random vectors of signs. Unfortunately, the

ZigZag body which we get approximates the Euclidean ball up to a (computable) universal constant, that is, we get a uniformly isomorphic approximation, and not an ε-isometric approximation. In fact, it will ε-approximate some other body, namely the dual of the floating body of the discrete cube (to be defined later), which is, however, *uniformly isomorphic* to the Euclidean ball (that is, there exists a universal constant such that the two bodies are isomorphic up to this constant). By this step we change from randomness of vectors on the sphere, which are hard to implement, to randomness connected with sign vectors, which are simpler to implement—each requires just n random bits. In this setting one needs a number of vectors which is linear in the dimension, and so, a number of bits which is quadratic in the dimension.

In the second step we will show that selecting only $Cn\log n$ random bits (which is the same as $C\log n$ random sign vectors) we are able to construct a ZigZag body uniformly isomorphic to Euclidean. For this we use a kind of derandomization procedure where we switch from random sign vectors to derandomized ones, in the expander framework of computer science. That is, in order to choose our N vectors we do not choose each one uniformly in the discrete cube. Instead, we put an expander of fixed degree on the vertices of the cube, then start at a random point on this expander and take a random walk of length N on the expander. Each point that we reach, we add to our collection of vertices.

We will show below that for $N = Cn\log n$ this construction will give us a ZigZag set isomorphic to the Euclidean ball (and, in fact, for large C depending on ε, this set will be ε-isometric to the dual of the floating body of the cube). The amount of randomness used, since this is implementing a random walk on a fixed degree expander (say, degree 3), is just $N = C'n\log n$ bits, which is like choosing $C'\log n$ random vectors.

The construction uses tools from expander theory, which we will explain in detail, and use again in subsequent sections. We start, however, with proving the desired property for the ZigZag body associated to *truly independent and random* points in the discrete cube.

3.1. The floating body of the discrete cube

The following theorem is the signed version of Theorem 4.

Theorem 5. *There exist universal constants C, c, c_0, c_1, C_2 such that for every dimension n, letting $N = [Cn]$, if $z_1, \dots, z_N$ are random points in $\{-1, 1\}^n$, then with probability greater than $1 - e^{-cn}$, the set*

$$\mathcal{K} = \left\{x \in \mathbb{R}^n\colon \exists i_1, \dots, i_{[2N/3]} \text{ with } \left|\langle x, z_{i_j}\rangle\right| < c_0\right\}$$

satisfies

$$c_1 D_n \subset \mathcal{K} \subset C_2 D_n.$$

We will not give the proof for this theorem, which is very similar to the proofs in [7], because derandomization of this theorem, although possible, is quite involved. Instead we prove a simpler theorem, which we later derandomize.

Theorem 5′ *(Weaker). There exist universal constants C, c, c_0, c_1, C_2 such that for every dimension n, letting $N = [Cn \ln n]$, if $z_1, \ldots, z_N$ are random points in $\{-1, 1\}^n$, then with probability greater than $1 - e^{-cn \log n}$, the set*

$$\mathcal{K} = \left\{ x \in \mathbb{R}^n \colon \exists i_1, \ldots, i_{[2N/3]} \text{ with } \left|\langle x, z_{i_j}\rangle\right| < c_0 \right\}$$

satisfies $c_1 D_n \subset \mathcal{K} \subset C_2 D_n$.

We remark that although this theorem is weaker than Theorem 5, the derandomization of this theorem and the derandomization of Theorem 5 both require (in the current state of affairs) the use of $C \log n$ random sign vectors (but from them one constructs a different number of final derandomized vectors in each case).

To prove this theorem we will need two simple estimates on the probabilities of entering slabs of a certain width when choosing a point uniformly in the discrete cube. The proof of the first uses the Berry–Esséen theorem, and we used the exact same lemma as in [7], see there for proof. The proof of the second inequality is a simple application of Markov's inequality.

Lemma 6. *There exists a universal constant $\alpha_0 > 0$ such that for every $\theta \in S^{n-1}$ we have*

$$\mathbb{P}\left[\left|\langle \varepsilon, \theta\rangle\right| \geqslant \alpha_0\right] \geqslant 1/2, \qquad \mathbb{P}\left[\left|\langle \varepsilon, \theta\rangle\right| \geqslant 2\right] \leqslant 1/4, \tag{1}$$

where $\varepsilon \in \{-1, 1\}^n$ is chosen uniformly.

We will also need the well-known Chernoff bound, namely that for Z_i independent random variables which take value 1 with probability p and value 0 with probability $1 - p$ the following holds:

(1) for every $\beta < p$ one has $\mathbb{P}[Z_1 + \cdots + Z_N \geqslant \beta N] \geqslant 1 - e^{-NI(\beta, p)}$,
(2) for every $\beta > p$ one has $\mathbb{P}[Z_1 + \cdots + Z_N > \beta N] \leqslant e^{-NI(\beta, p)}$, where

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

Proof of Theorem 5′. We build two nets, $\mathcal{N}_1$ a $(1/n)$-net on $c_1 S^{n-1}$ and $\mathcal{N}_2$ a $(1/n)$-net on $C_2 S^{n-1}$ (the constants c_1 and C_2 will be chosen later in the proof). It is well known that we can construct such nets with cardinalities smaller than $e^{n \ln(2c_1 n + 1)}$ and $e^{n \ln(2C_2 n + 1)}$, respectively. We define two bodies,

$$\mathcal{K}_1 = \left\{ x \in \mathbb{R}^n \colon \exists i_1, \ldots, i_{2N/3} \text{ with } \left|\langle x, z_{i_j}\rangle\right| < c_0 - \frac{1}{\sqrt{n}} \right\},$$

$$\mathcal{K}_2 = \left\{ x \in \mathbb{R}^n \colon \exists i_1, \ldots, i_{2N/3} \text{ with } \left|\langle x, z_{i_j}\rangle\right| < c_0 + \frac{1}{\sqrt{n}} \right\},$$

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 readily imply that $c_1 D_n \subset \mathcal{K} \subset C_2 D_n$, since for $y \in c_1 S^{n-1}$ there will be some $x \in \mathcal{N}_1$ with $|y-x| < 1/n$, and so if $|\langle x, z_i\rangle| < c_0 - 1/\sqrt{n}$ for some subset of indices i then for the same set of indices also $|\langle y, z_i\rangle| < c_0$. Similarly, every $y \in C_2 S^{n-1}$ will have an $x \in \mathcal{N}_2$ with $|y-x| < 1/n$, and for this x there will be at least $N/3$ indices i for which $|\langle x, z_i\rangle| \geqslant c_0 + 1/\sqrt{n}$. This implies that for y, for these indices, $|\langle y, z_i\rangle| \geqslant c_0$, and hence $y \notin \mathcal{K}$. Since $\mathcal{K}$ is star-shaped, this will complete the proof.

Let $x_j \in \mathcal{N}_1$. Then, since $|x_j| = c_1$, denoting $\theta = x_j/|x_j|$,

$$\mathbb{P}\left[z \in \{-1,1\}^n\colon \left|\langle z, x_j\rangle\right| \leqslant c_0 - \frac{1}{\sqrt{n}}\right] = \mathbb{P}\left[z\colon \left|\langle z, \theta\rangle\right| \leqslant \frac{c_0}{c_1} - \frac{1}{c_1\sqrt{n}}\right].$$

As long as, say, $c_1 \leqslant c_0/3$ (and n is large enough, larger than a universal constant) we have that this probability is greater than $3/4$, by Lemma 6.

We then use Chernoff's lemma, which says that for at least $2/3$ of the N experiments $\{|\langle z_i, x_j\rangle| \leqslant c_0 - 1/\sqrt{n}\}$ to succeed (x_j is fixed and z_i are random, $i = 1, \ldots, N$), when the probability of success is $3/4$, the chances are high, greater than

$$1 - e^{-NI(2/3,3/4)} = 1 - e^{-c_3 N}$$

for a universal

$$c_3 = \frac{2}{3}\ln\frac{8}{9} + \frac{1}{3}\ln\frac{4}{3}.$$

Therefore the chances that this will happen simultaneously for all $x_j \in \mathcal{N}_1$ are greater than $1 - e^{n\ln(c_1+2n)}e^{-c_3 N}$. For this to be greater than $1 - e^{-cn\log n}$ it suffices that

$$N > 2n\ln(c_1 + 2n)/c_3.$$

We now turn to the other side, which is surprisingly similar. Let $x_j \in \mathcal{N}_2$. Then, since $|x_j| = C_2$, denoting $\theta = x_j/|x_j|$,

$$\mathbb{P}\left[z \in \{-1,1\}^n\colon \left|\langle z, x_j\rangle\right| \leqslant c_0 + \frac{1}{\sqrt{n}}\right] = \mathbb{P}\left[z\colon \left|\langle z, \theta\rangle\right| \leqslant \frac{c_0}{C_2} + \frac{1}{C_2\sqrt{n}}\right].$$

As long as, say, $C_2 > 2c_0/\alpha_0$ (and n is large enough, larger than a universal constant) we have that this probability is smaller than $1/2$, by Lemma 6.

We then use Chernoff's lemma, which says that for at least $1/3$ of the N experiments $\{|\langle z_i, x_j\rangle| \geqslant c_0 + 1/\sqrt{n}\}$ to succeed (x_j is fixed and z_i are random, $i = 1, \ldots, N$), when the probability of success is $1/2$, the chances are high, greater than

$$1 - e^{-NI(1/3,1/2)} = 1 - e^{-c_4 N}$$

for a universal

$$c_4 = \frac{1}{3}\ln\frac{2}{3} + \frac{2}{3}\ln\frac{4}{3}.$$

Therefore the chances that this will happen simultaneously for all $x_j \in \mathcal{N}_1$ are greater than $1 - e^{n\ln(C_2+2n)}e^{-c_4 N}$. For this to be greater than $1 - e^{-cn\log n}$ it suffices that

$$N > 2n\ln(C_2 + 2n)/c_4. \qquad \square$$

3.2. Derandomization of the ZigZag construction

The main property of random independent points that we have used above is the following. Assume we are given a subset of the cube, $S \subset \{-1, 1\}^n$, whose size is a proportion p of the cube, $|S| = p2^n$. (For a finite set A we denote the number of points in A by $|A|$.) The probability that when choosing N random points in the cube, less than βN of them are in the subset S, for $\beta < p$, is exponentially small, according to Chernoff: $e^{-NI(\beta,p)}$. Similarly, the probability that one enters the subset S more than βN times, for $\beta > p$, is also exponentially small.

Instead of choosing the points randomly, we now choose them in the following way. Fix, once and forever, an expander of degree d on the cube (about explicitness of such an expander see Section 2). The structure of the expander has no relation with the structure of the cube. To choose N derandomized points $X_1, \ldots, X_N$ on the cube one starts with a random point on the cube, and then takes a random walk on the expander's edges, choosing at each step a random neighbor of the previously chosen vector, and adding to the collection each point stepped upon. After $N - 1$ steps one has a collection of N points on the cube, which we call "derandomized." The number of bits needed to collect these points is $n + (N - 1)[\log d]$.

We need to estimate how close to random these points are. More precisely, since the first point is uniformly distributed, so are all the rest. However, they are very strongly correlated, especially if the expander has a low degree. It turns out, however, that they are "random" enough for a Chernoff-type bound to hold. The first version of "derandomized Chernoff" that we give is from [13], but the key ideas go back to [1] and then [11,15]. The disadvantage of the following theorem is that it is relatively tight only in the case of sets S with probability not too close to 0 and not too close to 1. If one looks back at our proof of Theorem 5′, it is clear that we will not need more than this, since each time we use Chernoff there we use it for sets of probability either $3/4$ or $1/2$. However, when proving Theorem 5, and in particular the upper bound for the norm of a random sign-matrix, one uses sets of small probability, and then Theorem 7 does not suffice, and one has to dive further into the derandomization procedure and proofs to get meaningful estimates. In particular one has to use d which is not fixed but depends on n. We do this further on in this note, in Section 4, and apply it in various situations.

Theorem 7. [13] *There exists a universal constant c as follows. Given $n+2N$ random signs we can generate N vectors $Z_1, Z_2, \ldots, Z_N$ in $\{-1, 1\}^n$ such that for every $0 < p < 1$ and for any set $S \subset \{-1, 1\}^n$ with cardinality $p2^n$ the following holds: for every $\beta < p - 1/2N$ we have*

$$\mathbb{P}\big[\chi_S(Z_1) + \cdots + \chi_S(Z_N) \geqslant \beta N\big] \geqslant 1 - e^{-cN(p-\beta-1/(2N))^2}, \tag{2}$$

and for every $\beta > p + 1/2N$ *we have*

$$\mathbb{P}\big[\chi_S(Z_1) + \cdots + \chi_S(Z_N) > \beta N\big] \leqslant e^{-cN(\beta - p - 1/(2N))^2}. \tag{3}$$

Remark. The constant c can be brought as close to 1 as we wish, substituting $n + 2N$ by $n + CN$ for a large enough C.

The way $Z_1, Z_2, \ldots, Z_N$ in $\{-1, 1\}^n$ are generated is precisely through the random walk on an expander graph of degree 3. Copying the proof of Theorem $5'$ and using Theorem 7 instead of Chernoff's bound each time, we prove the following theorem.

Theorem 8. *There exist universal constants* C, c, c_0, c_1, C_2 *such that for every dimension* n*, given* $4C \ln n$ *random sign vectors* (*i.e.,* $4Cn \ln n$ *random signs*) *we can construct* $N = Cn \ln n$ *vectors* $z_1, \ldots, z_N \in \{-1, 1\}^n$ *such that with probability greater than* $1 - e^{-cn \log n}$*, the set*

$$\mathcal{K} = \big\{x \in \mathbb{R}^n \colon \exists i_1, \ldots, i_{[2N/3]} \text{ with } \big|\langle x, z_{i_j}\rangle\big| < c_0\big\}$$

satisfies $c_1 D_n \subset \mathcal{K} \subset C_2 D_n$.

4. More advanced derandomized Chernoff

In this section we describe and prove in detail a derandomized version of Chernoff's bound which can be used instead of Theorem 7, and is more flexible. So, for example, it can be manipulated to give meaningful results also when the probabilities involved are quite small (though we will have to pay for this by using a higher degree expander). We give a detailed proof which follows directly from the reasoning in the paper [2], and is intended for the convenience of the reader. We will use these estimates throughout the rest of the paper, and in particular in the next section where we derive Bernstein type inequalities.

Theorem 9. *Let* $S \subset \{-1, 1\}^n$ *be a set with* $|S| = p2^n$*. The probability that a random walk of* N *steps on an expander with degree* d *and expansion parameter* λ *on the cube* $\{-1, 1\}^n$*, starting from a uniformly random vertex, passes through the set* S *more than* k *times is at least*

$$\sum_{j=k}^{N} \binom{N}{j} (p - \lambda)^j (1 - p - \lambda)^{N-j},$$

and is at most

$$\sum_{j=k}^{N} \binom{N}{j} (p + \lambda)^j (1 - p + \lambda)^{N-j},$$

provided that $(1 - p)^2,\ p^2 \geqslant 2\lambda$.

Oppositely, we can bound the two quantities using the tails on the other side. Depending on the application, each of these estimates can be useful. That is, the same method proves that

Theorem 10. *Let $S \subset \{-1,1\}^n$ be a set with $|S| = p2^n$. The probability that a random walk of N steps on an expander with degree d and expansion parameter λ on the cube $\{-1,1\}^n$, with a uniformly random starting vertex, passes through the set S more than k times is at least*

$$1 - \sum_{j=1}^{k-1} \binom{N}{j} (p+\lambda)^j (1-p+\lambda)^{N-j},$$

and is at most

$$1 - \sum_{j=1}^{k-1} \binom{N}{j} (p-\lambda)^j (1-p-\lambda)^{N-j},$$

provided that $(1-p)^2, p^2 \geqslant 2\lambda$.

Remark 1. Here λ is the expansion parameter of the expander, that is, the second largest eigenvalue in absolute value of the normalized adjacency matrix, and so if we take an expander of degree d we can construct an explicit expander with $\lambda = c/\sqrt{d}$ (see, e.g., [4, 21,23]).

Remark 2. It is not important for the proof that the set of vertices of our expander is $\{-1,1\}^n$, and could be any other set.

Both theorems will follow from the following

Lemma 11. *In the notation and conditions of Theorems* 9 *and* 10*, the probability that the random walk passes through S* exactly *j times is at least*

$$\binom{N}{j} (p-\lambda)^j (1-p-\lambda)^{N-j}$$

and at most

$$\binom{N}{j} (p+\lambda)^j (1-p+\lambda)^{N-j}.$$

Proof *(following* [2]*).* Consider the 2^n-vector, indexed by the vertices of the cube, of the probabilities of being in the different vertices. Since we assume a uniform start, this vector to begin with is just the vector $e = (2^{-n}, \ldots, 2^{-n})$. At each step, we multiply the vector with the transition matrix of the expander, for which we know a bound on the second largest eigen-value.

To check whether our random walk passed through S *exactly* j times, we simply look at all sequences of S, S^c of length N which include S exactly j times. The probability that, for example, the walk starts at S^c, goes in the first step to S, then twice into S^c, and then into S again, is simply the sum of the coordinates of the vector $SAS^cAS^cASAS^ce$, where by abuse of notation we have denoted by S the matrix of the transformation which takes a vector into its restriction onto the subset of indices belonging to S, (thus $S_{i,i} = 1$ if $i \in S$ and 0, otherwise). We similarly defined the matrix S^c.

Fix a sequence of S and S^c, say $S_N S_{N-1} \cdots S_3 S_2 S_1$ where each S_i is either S or S^c. For convenience denote $v_1 = S_1 e$ and $v_i = S_i A v_{i-1}$. Our goal is to show that if the sequence contains S exactly j times (and so S^c exactly $N - j$ times) then the sum of the coordinates of v_N is less than $(p+\lambda)^j(1-p+\lambda)^{N-j}$ and greater than $(p-\lambda)^j(1-p-\lambda)^{N-j}$.

Notice that all the coordinates of v_i are positive so that we are in fact estimating its ℓ_1 norm $\|v_i\|_1$. It will complete the proof if we show that in case $S_{i+1} = S$ we have $(p-\lambda)\cdot\|v_i\|_1 \leqslant \|v_{i+1}\|_1 \leqslant (p+\lambda)\|v_i\|_1$, and in the case $S_{i+1} = S^c$ we have $(1-p-\lambda)\|v_i\|_1 \leqslant \|v_{i+1}\|_1 \leqslant (1-p+\lambda)\|v_i\|_1$.

We first decompose each vector v_i into a part x_i in direction e and a part y_i whose coordinates sum up to 0. Since x_i has all coordinates identical, and thus positive, and also the coordinates of v_i are positive, we see that $\|v_i\|_1 = \|x_i\|_1 = 2^{n/2}\|x_i\|_2$. Therefore it is enough to prove that for every i in case $S_{i+1} = S$ we have

$$(p-\lambda)\|x_i\|_2 \leqslant \|x_{i+1}\|_2 \leqslant (p+\lambda)\|x_i\|_2,$$

and in the case $S_{i+1} = S^c$ we have

$$(1-p-\lambda)\|x_i\|_2 \leqslant \|x_{i+1}\|_2 \leqslant (1-p+\lambda)\|x_i\|_2.$$

We do this by induction, adding to the induction hypothesis another one, namely that y_i satisfies for every i that

$$\|y_i\|_2 \leqslant \frac{1}{\sqrt{p(1-p)}}\|x_i\|_2.$$

Notice that for $i = 1$ this is satisfied. Indeed, in the case where $v_1 = Se$, that is, $(v_1)_\varepsilon = 2^{-n}$ if $\varepsilon \in S$ and 0, otherwise, we have that x_1 is the vector with all elements equal to $p2^{-n}$ and $(y_1)_\varepsilon = (1-p)2^{-n}$ for $\varepsilon \in S$ and $(y_1)_\varepsilon = -p2^{-n}$ for $\varepsilon \notin S$, so that

$$\|y_1\|_2 = \left(p2^n(1-p)^2 2^{-2n} + (1-p)2^n p^2 2^{-2n}\right)^{1/2} = \sqrt{(1-p)/p}\|x_1\|_2.$$

Similarly if $v_1 = S^c e$ we get the same result with p and $(1-p)$ interchanged. Thus, we have in both cases

$$\|y_1\|_2 \leqslant \frac{1}{\sqrt{p(1-p)}}\|x_1\|_2.$$

Assume by induction that we know the above for x_i, y_i. We want to prove it for the next stage. We begin with the case $S_{i+1} = S$. We have

$$v_{i+1} = SAv_i = Sx_i + SAy_i$$

since x_i is invariant under A. Same as the computation above for $v_1 = Sx_1$, we see that $Sx_i = px_i + z_i$ with $z_i \perp e$, the vector z_i having coordinates $(z_i)_\varepsilon = (1-p)(x_i)_\varepsilon$ when $\varepsilon \in S$ and $(z_i)_\varepsilon = -p(x_i)_\varepsilon$ when $\varepsilon \in S$ (recall that all $(x_i)_\varepsilon$ are equal). We decompose also $SAy_i = w_i + w_i'$ where $w_i' \perp e$ and w_i is in direction e. Thus, $x_{i+1} = px_i + w_i$ and $y_{i+1} = z_i + w_i'$. We first estimate $\|w_i\|_2$: since A has second eigen-value at most λ, and $y_i \perp e$, we have that the vector Ay_i, which is also perpendicular to e, has $\|Ay_i\|_2 \leqslant \lambda\|y_i\|_2$. It is not difficult to show that for $v \perp e$ the projection of Sv onto the span of e satisfies $\|P_e Sv\|_2 \leqslant \sqrt{p(1-p)}\|v\|_2$ (indeed, the worst case is when v is constant on S and another constant on S^c, which is similar to the case $i = 1$ which we studied above). Using this for $v = Ay_i$ we get that $\|w_i\|_2 \leqslant \lambda\sqrt{p(1-p)}\|y_i\|_2$. From the induction hypothesis we have that $\|w_i\|_2 \leqslant \lambda\|x_i\|_2$. Using the triangle inequality (in both directions) we get the desired property of $\|x_i\|_2$. We turn to y_{i+1}, for which we see

$$\|y_{i+1}\|_2 \leqslant \|z_i\|_2 + \|w_i'\|_2 \leqslant \sqrt{p(1-p)}\|x_i\|_2 + \|Ay_i\|_2 \leqslant \sqrt{p(1-p)}\|x_i\|_2 + \lambda\|y_i\|_2.$$

By the induction hypothesis we get

$$\|y_{i+1}\|_2 \leqslant \sqrt{p(1-p)}\|x_i\|_2 + \frac{\lambda}{\sqrt{p(1-p)}}\|x_i\|_2$$

and using what we already showed, namely $\|x_{i+1}\|_2 \geqslant (p-\lambda)\|x_i\|_2$, together with the assumption $\lambda \leqslant p^2/2$, we get that

$$\|y_{i+1}\|_2 < \frac{1}{\sqrt{p(1-p)}}\|x_{i+1}\|_2.$$

We now have to repeat this in the case $S_i = S^c$, but everything remains unchanged, since the expressions are symmetric with respect to p and $1-p$, only that now we have to use the condition $\lambda \leqslant (1-p)^2/2$. This completes the proof by induction. □

Remark. Of course, one can analyze these estimates, in a Chernoff-type way. For example, if $k > (p+\lambda)N$ then

$$\begin{aligned}\mathbb{P}[\text{more than } k \text{ times in } S] &\leqslant \sum_{j=k}^{N} \binom{N}{j} (p+\lambda)^j (1-p+\lambda)^{N-j} \\ &\leqslant \left(\frac{1-p+\lambda}{1-p-\lambda}\right)^{N-k} \sum_{j=k}^{N} \binom{N}{j} (p+\lambda)^j (1-p-\lambda)^{N-j}\end{aligned}$$

$$= \left(1 + \frac{2\lambda}{1-p-\lambda}\right)^{N-k} \sum_{j=k}^{N} \binom{N}{j} (p+\lambda)^j (1-p-\lambda)^{N-j}$$

$$\leqslant \left(1 + \frac{2\lambda}{1-p-\lambda}\right)^{N-k} e^{-NI(k/N, p+\lambda)}.$$

For $\lambda = 1/N$ we get that as long as $p > 2/\sqrt{N}$ and $k \leqslant pN + 1$, the probability of more than k successes is less than

$$e^2 \cdot e^{-NI(k/N, p+1/N)} \leqslant e^2 (p + 1/N)^k e^{-Nu(k/N)}$$

(where $u(\beta) = \beta \ln \beta + (1-\beta)\ln(1-\beta)$). For probabilities not very small this is already a good bound.

Sometimes this is not sufficient. If we take, as we do in some of the applications, $\lambda = e^{-2\sqrt{N}}/2$ we get that as long as $k \geqslant pN + 1$ and when $p \geqslant e^{-\sqrt{N}}$ the probability of more than k successes is less than

$$e^2 \cdot e^{-NI(k/N, p+\lambda)} \leqslant e^2 (2p)^k e^{-Nu(k/N)}.$$

5. Derandomized Bernstein-type inequalities

To prove upper bounds in asymptotic geometric analysis, one of the standard tools is Bernstein's inequality. This inequality bounds in a strong way the deviation of the average of independent random variables from their average mean, based on their individual tail estimates. In this section we would like to get similar estimates for the derandomized random variables. Put it another way, given not many signs we produce N vectors of signs in the following way (as in the previous section): put on $\{-1, 1\}^n$ an expander of degree d with expansion parameter $\lambda \approx 1/\sqrt{d}$. Consider a random walk which starts at a random point on the cube, of length N on this expander. It requires only $n + N \log d$ bits. Call the acquired points $X_1, \ldots, X_N$. We want to show that they satisfy a Bernstein-type inequality in some appropriate sense. We first show the following theorem, which is rather weak unless the function is well bounded in advance.

Theorem 12. *The N vectors $X_1, X_2, \ldots, X_N$ in $\{-1, 1\}^n$ described above satisfy that if for some function $f : \{-1, 1\}^n \to \mathbb{R}$, for every $t > t_0$*

$$\left|\left\{X \in \{-1, 1\}^n : \ f(X) > t\right\}\right| \leqslant 2^n e^{-Kt^2},$$

then we have for every $s > \max(t_0, C_1/\sqrt{K}, C_2\sqrt{\lambda} b)$

$$\mathbb{P}\left[\sum_{i=1}^{N} f(X_i) > CsN\right] \leqslant e^{-c_1 K s^2 N} + e^{2\lambda N} e^{-c_1 s N \ln(1/8\lambda)/b} e^{-Nu(c_1 s/b)},$$

where c_1, C_1, C_2 and C are universal constants, and b is the upper bound of f (which from the assumption is always less than $\sqrt{n/(K\ln 2)}$). In fact, c_1 can be as large as we wish, and then we take $C = 4 + c_1$.

The term $u(\cdot)$ is at least $-\ln 2$ so in the case of $cs\ln(d/8) > b$ it will be absorbed in the other term. In other cases we can make it small using the fact that u tends to 0 as its argument tends to 0, and that the nontrivial case is when s is much smaller than b.

Proof. First we remark that the condition in the theorem promises that $f(X) < \sqrt{n/K\ln 2}$ for every X. So we have an upper bound on f, which we denote by $b \leqslant \sqrt{n/K\ln 2}$. For $j = \log s + 1, \log s + 2, \ldots, \log(\sqrt{\ln(1/2\lambda)/2K}) =: j_\infty$ we define

$$A_j = \left\{X \in \{-1,1\}^n\colon\ 2^{j-1} < f(X) \leqslant 2^j\right\},$$

so that by our assumption $\mathbb{P}[X_i \in A_j] \leqslant e^{-K2^{2j-2}}$ (where we have used the assumption $s > t_0$). We also define

$$A_\infty = \left\{X \in \{-1,1\}^n\colon\ 2^{j_\infty} < f(X)\right\}.$$

Thus, $\mathbb{P}[X_i \in A_\infty] \leqslant \sqrt{2\lambda}$.

We set $m_j = Ns2^{-j}/(j - \log s)^2$, and $m_\infty = c_1 sN/b$. We measure the probability of the following event: out of the N variables X_i, for every j, no more than m_j of them are in A_j. This event is included in the event that

$$\frac{1}{N}\sum_{i=1}^{N} f(X_i) \leqslant s\left(1 + \sum_{j=1}^{j_\infty}\frac{1}{j^2} + c_1\right) \leqslant Cs.$$

We will estimate the probability of the complementary event. It is less than the sum over j over the individual probabilities

$$P_j = \mathbb{P}[\text{more than } m_j \text{ of the } X_i\text{'s are in } A_j].$$

We start with A_∞. The number of successes we ask for is $c_1 sN/b$, and we want to show that the probability for this is small. Indeed, the probability of each success is less than $\sqrt{2\lambda}$, and by the derandomized Chernoff we get that the probability for at least $k = c_1 sN/b$ (which is greater than $\sqrt{8\lambda}N + 1$ since we assumed $s > C_2\sqrt{\lambda}b$) successes is less than

$$e^{\lambda N}\mathbb{P}[\text{more than } c_1 sN/b \text{ successes when } p = \sqrt{8\lambda}] \leqslant e^{\lambda N}e^{-csN\ln(1/8\lambda)/b}e^{-Nu(c_1 s/b)}.$$

We continue with the other sets, namely estimate P_j for $j = \log s + 1, \ldots, j_\infty$. As long as

$$s2^{-j}/(j - \log s)^2 > e^{-K2^{2j-2}} + \lambda \tag{4}$$

(which is satisfied by the condition on s, namely the lower bound on s in terms of K, notice that the term λ does not interfere because we made sure that $p = e^{-K2^{2j-2}}$ is larger than $\sqrt{2\lambda}$), this probability is small, and by the derandomized Chernoff it is smaller than

$$\left(e^{-K2^{2j}} + \lambda\right)^{Ns2^{-j}/(j-\log s)^2} \cdot e^{-Nu(s2^{-j}/(j-\log s)^2)}.$$

Since the λ term is harmless here, and also the term $u(\cdot)$ can be absorbed in the other term, we can bound this by

$$\left(e^{-cK2^{2j}}\right)^{Ns2^{-j}/(j-\log s)^2}.$$

The sum of these probabilities converges and is equivalent to the first term, which is what we wanted. □

By using $\lambda \approx e^{-\sqrt{K}\sqrt{n}}$ we see the following.

Corollary 13. *We can use $C_0 n^{1/2} N$ bits to create N vectors $X_1, X_2, \ldots, X_N$ in $\{-1, 1\}^n$ satisfying that if for some function $f : \{-1, 1\}^n \to \mathbb{R}$, for every $t > t_0$*

$$\left|\left\{X \in \{-1, 1\}^n : \ f(X) > t\right\}\right| \leqslant 2^n e^{-Kt^2},$$

then we have for every $s > \max(t_0, c/\sqrt{K})$

$$\mathbb{P}\left[\sum_{i=1}^{N} f(X_i) > CsN\right] \leqslant e^{-c\sqrt{K}sN},$$

where c, C_0 and C are universal constants.

The result above already gives a significant reduction in randomness, which we will use in some of our geometric consequences, but not yet to the level we aimed at. Indeed, ideally we would hope to be able to use a constant degree expander, same as we used in the ZigZag construction in Section 3.2. However, because we had to deal with very small probabilities, our expander had to be of such a huge degree that we only decreased the number of vectors in any application which will use to above estimate from nN to $n^{1/2}N$. Usually in applications we will not have one function but a net of functions, in which case (for, say, K and s fixed) N must be of the order of n, so the decrease in randomness is from n^2 to $n^{3/2}$.

However, in some applications a weaker statement than Corollary 13 is enough, in which we have a Bernstein-type theorem for a subset of the vectors. For example, for proofs regarding sections of the cross polytope which we discussed in the introduction (and appear as Theorem 15 in the next section) we do not need the full strength of a Bernstein-type estimate to hold on all vectors. What we can make do of is a variation on Theorem 12 of the following form.

Theorem 14. *The N vectors $X_1, X_2, \ldots, X_N$ which are attained by taking a random walk from a uniform starting vertex on an expander with vertices $\{-1,1\}^n$, of degree d and with expansion parameter λ, satisfy that if for some function $f:\{-1,1\}^n \to \mathbb{R}$, for every $t > t_0$*

$$\left|\left\{X \in \{-1,1\}^n:\ f(X) > t\right\}\right| \leqslant 2^n e^{-Kt^2},$$

then we have for every $N > k > NC'/(\log(1/2\lambda))$ and every $s > \max(t_0, C_1/\sqrt{K})$ that

$$\mathbb{P}\left[\exists I,\ |I| = N-k,\ \sum_{i\in I} f(X_i) < CsN\right] \geqslant 1 - \left(e^{-cKs^2N} + e^{-c'N}\right), \tag{5}$$

where C', c', C_1, C_2 and C are universal constants. In fact c' can be as large as we wish, taking $C' = 2\log(e^{c'+3})$.

Proof. We sketch the proof, elaborating only on the parts which are different from the proof of Theorem 12. Define A_j for $j = \log s + 1, \log s + 2, \ldots, j_\infty$ and A_∞ as before, however, this time let

$$j_\infty = \frac{1}{2}\log\left(C'N/(2Kk)\right) \quad \text{where } C' = 2\log\left(e^{c'+3}\right).$$

Thus,

$$\mathbb{P}[X_i \in A_\infty] \leqslant e^{-K2^{2j_\infty}} = 2^{-C'N/2k}\ (> \sqrt{2\lambda}), \quad \text{and}$$
$$\mathbb{P}[X_i \in A_j] \leqslant e^{-K2^{2j-2}}.$$

Set again $m_j = Ns2^{-j}/(j - \log s)^2$, and this time $m_\infty = k$. We measure the probability of the event that out of the N variables X_i, for every j, no more than m_j of them are in A_j. This event is included in the event that there exists $I \subset [1,\ldots,N]$ with $|I| = N-k$ and

$$\frac{1}{N}\sum_{i\in I} f(X_i) \leqslant s\left(1 + \sum_{j=1}^{j_\infty} \frac{1}{j^2}\right) \leqslant Cs,$$

where, of course, the subset I is chosen to be all indices not in A_∞ (and if there are too many one can just drop the extra amount).

The probability of the complementary event is less than the sum over j over the individual probabilities P_j that more than m_j of the X_i's are in A_j. The estimate for $j = \log s + 1, \ldots, j_\infty$ is the same as before. Regarding A_∞, the number of successes we ask for is k, which is greater than $\lambda N + 1$ by the assumption, and even more so if we take larger C'. The probability of each success is less than $2^{-C'N/2k}$, and by the derandomized Chernoff we get that the probability for at least k successes is less than

$$e^{\lambda N}\mathbb{P}\left[\text{more than } k \text{ successes when } p = 2^{-C'N/2k+1}\right] \leqslant e^{-c'N}.$$

Putting all these probabilities together, the proof is complete. □

6. Sections of ℓ_1^n

We return to the example mentioned in the introduction, of sections of the cross-polytope, that is, the unit ball of ℓ_1^n. The fact that the cross polytope $B(\ell_1^n)$ has sections of proportional dimension which are isomorphic to Euclidean has been known for long, see [12] for sections of small proportional dimension, in which case you get almost isometric embeddings, and [18] for $(n/2)$-dimensional sections isomorphic to Euclidean up to universal constants. The proofs show that a *random* section satisfies this with high probability. Here randomness is with respect to the Haar measure on the Grassmanian of all, say, $(n/2)$-dimensional subspaces. In the later work of Schechtman (see [9,34]) it was shown that in fact the image of a random matrix of signs provides such a section as well. This is a first step toward explicitness, since for finding an $(n/2)$-dimensional subspace with this property there exist "only" $2^{n^2/2}$ matrices of signs to search among. A further step was provided recently in [35] and described also in [6] where it was shown that by choosing the $m \times m$ identity matrix, or any fixed orthogonal matrix, suitably normalized, and adding just δm rows of random signs, one can generate a section of $\ell_1^{m+\delta m}$ which is isomorphic to Euclidean up to a constant depending only on δ. Thus, denoting $n = (1+\delta)m$, we in fact get a section of dimension $\frac{1}{1+\delta}n$, and, moreover, the constants of isomorphism depend only polynomially on δ. So, one can use little randomness and get even very high-dimensional sections of ℓ_1^n which are isomorphic to Euclidean.

In this section we show that one can use even less randomness, and in fact just $C\log n$ random sign vectors can be used to construct, with some additional explicit steps, a section which is uniformly isomorphic to Euclidean. What we show is that the kernel of a certain matrix which we build using $C\log n$ random sign-vectors, is a subspace which gives a good, that is, close to Euclidean, proportional section. To build a subspace which is *spanned* by sign-vectors and is a good section we need to use more randomness, namely $C\sqrt{n}$ random vectors, with which, in an explicit way, we generate the section.

6.1. The Kernel method

In this section we prove the following theorem.

Theorem 15. *For any $0 < \varepsilon < 1$ we can use $C_0 \log n$ random sign vectors (i.e., $C_0 n \log n$ random bits) to create $N = \varepsilon n$ vectors $\xi_1, \xi_2, \ldots, \xi_N$ in $\{-1,1\}^n$ satisfying with probability greater than $1 - e^{-cN}$ that for every $y \in \mathbb{R}^n$ with $\langle y, \xi_i \rangle = 0$ for every i, we have*

$$c_1(\varepsilon)|y| \leqslant \frac{1}{\sqrt{n}} \sum_{j=1}^{n} |y_i| \leqslant |y|,$$

where $c(\varepsilon)$ depends polynomially on ε, and c, C_0 are universal.

Proof. The upper bound is true for any $y \in \mathbb{R}^n$. For the lower bound we cover $\sqrt{n}B(\ell_1^n)$ by $M = e^{\delta n}$ balls each of radius $r = c_1(1/\delta \ln(1/\delta))^{1/2}$, which is possible by the covering estimates of Schütt [36].

The parameter δ will be chosen later to be linear in ε. We then look only at the centers z_i of balls which satisfy $|z_i| > R$ where R will be chosen later to be of order r. The r-balls around these points thus cover $\sqrt{n}B(\ell_1^n) \setminus (R+r)D_n$.

We will generate the vectors $\xi_1, \ldots, \xi_N$ by walking N steps on an expander of degree n^{α_0} (for some constant α_0 we choose later, and which depends only on ε; a careful analysis will show it should be chosen proportional to ε^{-1}) whose vertices are $\{-1, 1\}^n$. Let $\mathcal{N}$ be a $(1/2)$-net on the sphere, with $|\mathcal{N}| \leqslant e^{n \ln 5}$.

We claim that with high probability the following two things happen simultaneously:

(a) for each $i = 1, \ldots, M$ we have that there is a subset A_i of at least $N/3$ indices j such that $c_2 R \leqslant |\langle \xi_j, z_i \rangle|$ for every $j \in A_i$ (and c_2 is universal);
(b) for every $w \in \mathcal{N}$ there is a subset $A'_w \subset [1, \ldots, N]$ with $|A'_w| = N - \alpha_1 N/\log n$ and we have $\sum_{j \in A'_w} |\langle w, \xi_j \rangle| \leqslant N C_3 |w|$ for every $w \in \mathcal{N}$ (α_1 and C_3 will be chosen later, depending only on ε).

We first show that the two properties hold and then explain why this proves the theorem.

Property (a) follows as before by using Chernoff with fixed probability. Indeed, for any z_i we have $\mathbb{P}[|\langle \xi_j, z_i \rangle| \geqslant R c_2] \geqslant 1/2$ for a universal c_2 (see, e.g., [7]). Therefore with exponential probability in N, for a random walk on an expander starting at a uniform vertex of $\{-1, 1\}^n$ we will have with probability greater than $1 - e^{-c'N}$ that at least $1/3$ of the indices j satisfy the above (where c' is universal). The fact that out expander has very high degree only improves the constants involved in this estimate. Since δ can be chosen as a small proportion of ε, this probability is sufficient to take care simultaneously on all the z_i.

Part (b) follows from Theorem 14. We use it with $d = n^{\alpha_0}$, $\lambda = n^{-\alpha_0/2}$ and $k = \alpha_1 N/\log n$ (for some universal α_1 to be chosen later). Notice that the condition of Theorem 14, namely $k > 2NC'/(\alpha_0 \log 2n)$, is satisfied if $\alpha_1 > 3C'/\alpha_0$. The net we have to take care of has cardinality $\leqslant e^{n \ln 5}$, so we need to make sure the probability in (5) is small enough so that we can take care of all points in the net. The function, as before, will be $f_w(X) = |\langle w, X \rangle|$ so that we know

$$\left| \left\{ X \in \{-1, 1\}^n \colon\ f_w(X) > t \right\} \right| \leqslant 2^n e^{-t^2/4},$$

for every $t > 2$, in particular K and t_0 from Theorem 14 are universal.

Then for a fixed $w \in \mathcal{N}$ we have that for some universal constant C_4

$$\mathbb{P}\left[\exists I,\ |I| = N - \alpha_1 N/\log n,\ \sum_{i \in I} f(X_i) < C_3 N \right] \geqslant 1 - e^{-C_4^2 N} + e^{-c'N}.$$

We can make C_4 and c' large enough by picking large C', C_3. Recall that $N = \varepsilon n$ so we need to have a large factor to be able to take care of the whole net. This is what gives $\alpha_0 \simeq \varepsilon^{-1}$ and $C_3 \simeq \varepsilon^{-1/2}$. With these choices we can do this simultaneously for all w's in $\mathcal{N}$. We have thus shown that also (b) holds with high probability, at least $1 - e^{-c''n}$, provided $\alpha_1 > 2C'/\alpha_0$.

The two properties imply the theorem. We first use (b) to deduce something weaker which holds not only on the net but over the whole sphere: an upper bound on a proportion

of indices, and this is enough to complete the proof. If $y \in \mathrm{Ker}(\xi_j)$ then there is a closest point z_i to y in the net. This i produces a subset A_i of $N/3$ indices such that for every $j \in A_i$ we know $|\langle z_i, \xi_j\rangle| \geqslant c_2 R$. We consider $w = y - z_i$. We can successively approximate it (actually, its normalization $w' = w/|w|$) by points in the net which we took care of in (b). We do this only up to a finite number, that is, we write

$$w' = w_0 + \theta_1 w_1 + \theta_2 w_2 + \cdots + \theta_{p-1} w_{p-1} + \theta_p r_p$$

with $|w_j| = |r_p| = 1$, and $|\theta_j| \leqslant (1/2)^j$, and each w_j is in the net. We do this up to $p = \alpha_2 \log n$ so that $|\theta_p r_p| < C/\sqrt{n}$, where we have taken, say, $\alpha_2 = 1/(2\log 2)$. For each w_j, for $j = 1, \ldots, p$, there corresponds I_j of cardinality $N - \alpha_1 N/\log n$. Provided $\alpha_1\alpha_2 \leqslant 1/4$, their intersection $\bigcap_{j=1}^{p-1} I_j$ is of cardinality at least $3N/4$. We denote the intersection $\bigcap_{j=1}^{p-1} I_j \cap A$ by A'', and we thus have $|A''| > N/12$.

We now see that

$$\frac{1}{N}\sum_{i\in A''} \big|\langle w', \xi_i\rangle\big| \leqslant \frac{1}{N}\sum_{j=0}^{p-1}\left(\frac{1}{2}\right)^j \sum_{i\in A''} \big|\langle w_j, \xi_i\rangle\big| + C \leqslant 2C_3 + C =: C_3',$$

and this inequality is homogeneous so we have

$$\frac{1}{N}\sum_{i\in A''} \big|\langle w, \xi_i\rangle\big| \leqslant C_3'|w| \tag{6}$$

(note $C_3' \approx 1/\sqrt{\varepsilon}$).

In other words, we were able this way to transfer an upper bound to the whole sphere. We complete the proof by writing

$$|y - z_i| \geqslant \frac{1}{C_3' N}\sum_{j\in A''} \big|\langle y - z_i, \xi_j\rangle\big| = \frac{1}{C_3' N}\sum_{j\in A''} \big|\langle z_i, \xi_j\rangle\big| \geqslant \frac{c_2 R}{12 C_3'}.$$

For $R = C_5 r$ with C_5 big enough, We see that $|y - z_i| > r$ so that the kernel does not intersect the balls covering $\sqrt{n}B(\ell_1^n)$ above level $R + r$, and the proof is complete.

The number of random bits used is $Cn\log n$. This does not depend on ε, however the isomorphism constant does depend, polynomially, on ε. □

In fact, inside this proof by (b) which then implied (6) we have shown the following corollary of Theorem 14 (notice that in this corollary only $n + 2N\alpha\log n$ random bits are used).

Corollary 16. *There exists a universal c and for any $\beta < 1$ there exists constants $\alpha(\varepsilon)$, $C(\varepsilon)$ depending only on β and on $\varepsilon = (N/n)$ such that the N vectors $X_1, X_2, \ldots, X_N$ which are attained by taking a random walk from a uniform starting vertex on an expander with*

vertices $\{-1,1\}^n$*, with expansion parameter* $n^{-\alpha(\varepsilon)}$*, satisfy that with probability* $1-e^{-cn}$ *one has for any* $w\in\mathbb{R}^n$ *that there is an* $I_w\subset[1,\dots,N]$ *with* $|I_w|=\beta N$ *and*

$$\sum_{i\in I_w}\big|\langle X_i,w\rangle\big|<C(\varepsilon)N. \tag{7}$$

In fact, since $C(\varepsilon)=C/\sqrt{\varepsilon}$*, we have for a universal* C_0 *that with probability greater than* $1-e^{-cn}$

$$\frac{1}{N}\sum_{i\in I_w}\big|\langle X_i,w\rangle\big|<C_0\sqrt{\frac{n}{N}}. \tag{8}$$

6.2. The Image method

The "dual" way of creating a section which is isomorphic to Euclidean is to take the image of a random sign-matrix instead of its kernel. One would hope that the same derandomization technique would work. However, we were not able to implement it here, and can only show the weaker statement where the vectors are chosen using an expander of degree $e^{\sqrt{n}}$, that is, each vector require order $\sqrt{n}$ of random bits. Choosing N sign vectors in this way, we can show that the subspace of R^N which is the image of the matrix whose rows are these N vectors, is a subspace in which the ℓ_1^N norm is equivalent to a Euclidean norm. We prove

Theorem 17. *We can use* $C_0\sqrt{n}$ *random sign-vectors in* $\{-1,1\}^n$ *(i.e.,* $C_0n^{3/2}$ *bits) to construct* $N=C_1n$ *sign-vectors* $\xi_1,\xi_2,\dots,\xi_N$ *in* $\{-1,1\}^n$ *satisfying that for every* $y\in\mathbb{R}^n$ *we have*

$$c_2|y|\leqslant\frac{1}{N}\sum_{i=1}^{N}\big|\langle\xi_i,y\rangle\big|\leqslant C_3|y|,$$

where C_0,C_1,c_2 *and* C_3 *are universal.*

Proof. We begin with the upper bound. It follows from Corollary 13, since we have created the vectors as indicated in the corollary. More precisely, we define for $|w|=1$ functions $f_w:\{-1,1\}^n\to\mathbb{R}$ by $f_w(X)=|\langle w,X\rangle|$ so that we know

$$\big|\big\{X\in\{-1,1\}^n\colon\ f_w(X)>t\big\}\big|\leqslant 2^ne^{-t^2/4},$$

for every $t>2$ (see, for example, [5]). Then for a fixed $w\in S^{n-1}$ we have that for some constant C (depending only on C_1)

$$\mathbb{P}\left[\sum_{i=1}^{N}\big|\langle w,X_i\rangle\big|>CN\right]\leqslant e^{-N(2\ln 5/C_1)}=e^{-2n\ln 5}.$$

This means we can do this simultaneously for all w's in a $(1/2)$-net on the sphere, then by successive approximation we have that the inequality holds for all points $w \in S^{n-1}$, and by homogeneity of the inequality we get the upper bound.

For the lower bound we pick a δ-net on the sphere S^{n-1} and for every w in this net show that with high probability there are at least $N/3$ indices for which $|\langle \xi_i, y\rangle| > 3c_2$, where c_2 is universal. This is true by the derandomized Chernoff for fixed probability, and requires that the exponent in N we get for the probability is greater than the number of points in the net, which holds if we assume $N > C_1(\delta)n$. We pick $\delta = c_2/2C_3$ so that having a lower bound on the net we may transfer it to a lower bound on the whole sphere. The proof is thus complete. □

Remark. We can use the "Image method" but instead of all of the vectors being random, or pseudo-random, we may take the first n vectors be the standard Walsh basis vectors, or any other orthogonal matrix suitable normalized. We then add to these vectors εn other vectors which we generate using an expander. The advantage of this method is that we automatically have a lower bound for "most" of the points. Thus the net which we had to use in the lower bound proof in the above theorem can be taken in this scheme to be of a much smaller size (as small as we want, depending on the lower bound we wish to obtain). This is why we can make do with only ε more vectors. The lower bound we get is polynomial in terms of ε. Notice, however, that although the lower bound goes smoothly (and we could get it with vectors generated by a degree three expander), the upper bound is as difficult as above and we need the very high degree expander to ensure it.

If we want a statement which is of image-type, and use the same trick of avoiding Bernstein's bound on the whole set of vectors, thus using less randomness we can prove the following theorem. Its proof is in similar lines to the proof in the section on the Kernel method, and we omit it.

Proposition 18. *We can use $C_0 n \log n$ bits to create $N = C_1 n$ vectors $\xi_1, \xi_2, \ldots, \xi_N$ in $\{-1, 1\}^n$ satisfying that for every $y \in \mathbb{R}^n$ we have a subset $I_y \subset [1, \ldots, N]$ with $|I_y| \geqslant N/4$ and*

$$c_2|y| \leqslant \frac{1}{N}\sum_{i\in I_y}\left|\langle \xi_i, y\rangle\right| \leqslant C_3|y|,$$

where C_0, C_1, c_2 and C_3 are universal.

7. Sections of bodies with finite volume ratio

Using the Kernel method as in Section 6.1 one notices that in fact we have not used any special properties of $B(\ell_1^n)$, apart from a fact regarding its covering numbers. The

body $\sqrt{n}B(\ell_1^n)$, which contains D_n, may be covered by an exponential number of copies of rD_n, where the constant in the exponent decreases with r, that is

$$N\left(\sqrt{n}B\left(\ell_1^n\right), rD\right) \leqslant e^{nf(r)}$$

with $f(r) \to 0$ as $r \to \infty$. Thus, any body possessing this property will satisfy the conclusions valid for $\sqrt{n}B(\ell_1^n)$. Moreover, we did not have to cover all of $\sqrt{n}B(\ell_1^n)$, but just $\sqrt{n}B(\ell_1^n) \cap RD$ where r was a proportion of R.

It was shown in [19] that the class of bodies with finite volume ratio satisfies this type of inequality. More precisely, Proposition 8 from [19] states that if $D \subset K \subset R^n$ and $A = (\mathrm{Vol}(K)/\mathrm{Vol}(D))^{1/n}$ is the volume ratio, then we have for every $R > 1/\sqrt{\ln(2A)}$ that

$$\frac{M^*(K \cap RD)}{R} \leqslant f(R, A),$$

where $f(R, A) \to 0$ as $R \to \infty$.

Using Sudakov this implies that for such K and every R one has

$$N(K \cap RD, rD) \leqslant \exp\left(n(R/r)^2 C^2 f(R, A)\right).$$

In particular, if the ratio R/r is fixed and $R \to \infty$ this number can be reduced to be as small an exponent as desired, by enlarging R.

Therefore we can repeat the proof of Theorem 15. We arrive at the following:

Theorem 19. *For any body K with finite volume ratio A, setting it in the position where the maximal volume ellipsoid of K is D_n, and any orthogonal basis, for any $\varepsilon > 0$ we can use $C_0(A)\log n$ random sign vectors (i.e., $C_0(A)n\log n$ random bits) to create $N = \varepsilon n$ vectors $\xi_1, \xi_2, \ldots, \xi_N$ in $\{-1, 1\}^n$ satisfying that for any $y \in \mathbb{R}^n$ with $\langle y, \xi_i\rangle = 0$ for every i, we have*

$$c_1(\varepsilon, A)|y| \leqslant \|y\|_K \leqslant |y|.$$

Note that this is a derandomized version of a theorem which also holds, of course, when the N sign-vectors are selected uniformly at random. In this form it was proved in the paper [19].

8. Low M^* estimate

In this section we derandomize yet another well-known theorem from asymptotic convex geometry, namely Milman's low M^* estimate, which was originally proved as part of the proof of his QS-theorem, see [25] (where it is not explicitly stated) and [24]. The estimate states that a random section of co-dimension k of a symmetric convex body K has radius at most $M^* f(k/n)$. The estimate in [24] was $f(\lambda) = C/\lambda$ for a universal C. This estimate was then improved in [31] to $C/\sqrt{\lambda}$ and the best estimate as $\lambda \to 0$ is due to Gordon [14], and is $f(\lambda) = (1 + O(1))/\sqrt{\lambda}$. All these results correspond to random

sections with respect to the Haar measure on $G_{n,n-k}$. Milman and Pajor in [28] showed that a low M^* estimate also holds in the case of a random subspace chosen as the kernel of a random sign matrix (we call this "signed low M^*"). They showed, for this choice of random subspace, that a low M^* estimate holds with $f(\lambda) = C/\lambda$, and remarked that $C/\sqrt{\lambda}$ also follows from their method, adding the result in [5].

Below we give a derandomized version of this estimate, and recover their linear dependence $f(\lambda) = C/\lambda$ in the derandomized case. For this we will use Corollary 16.

Theorem 20. *Let K be a convex body in $\mathbb{R}^n$ equipped with the standard Euclidean structure. Let $k < n$. For some $\alpha = \alpha(k/n)$ we can use $\alpha \log n$ random sign-vectors to generate explicitly sign-vectors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k \in \{-1, 1\}^n$, so that with probability $1 - e^{-ck}$ one has for $E = \bigcap_{i=1}^k \mathrm{Ker}(\varepsilon_i)$ that*

$$K \cap E \subset C_1 \frac{M^*(K)}{(k/n)} D_n,$$

where C_1 is universal.

Remark 1. As before, the vectors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k$ are attained by taking a random walk from a uniform starting vertex on an expander with vertices $\{-1, 1\}^n$, with expansion parameter $n^{-\alpha}$.

Remark 2. The same conclusion of course holds when we build the vectors using a different orthogonal basis.

Proof. We use Sudakov inequality to cover K with $T = e^{c_1 k}$ balls of radius $r = 2c_1^{-1/2} M^* \sqrt{n/k}$. Call their centers $x_1, \ldots, x_T$. As before, by standard Chernoff for derandomized vectors we can make sure that as long as c_1 is small enough, for every j we have a subset $|I_j| = k/3$ of indices such that for $i \in I_j$ we have $|\langle x_j, \varepsilon_i \rangle| \geqslant c_2 |x_j|$. We can also have, at the same time, that the conclusion (8) of Corollary 16 holds, with $\beta = 3/4$, $N = k$ and $X_i = \varepsilon_i$. Thus, for any $y \in E$ we have

$$|y - x_j| \geqslant \frac{1}{C_0\sqrt{nk}} \sum_{i \in I_w} \left|\langle y - x_j, \varepsilon_i \rangle\right| = \frac{1}{C_0\sqrt{kn}} \sum_{i \in I_w} \left|\langle x_j, \varepsilon_i \rangle\right|$$

$$\geqslant \frac{1}{C_0\sqrt{kn}} \sum_{i \in I_w \cap I_j} \left|\langle x_j, \varepsilon_i \rangle\right| \geqslant \frac{\sqrt{k/n}}{12 C_0} c_2 |x_j|.$$

If we look only at $|x_j| > R$ then the union of r-balls around such centers covers the set $K \setminus (R + r) D_n$, and so we get that for $R = 12 C_0 r \sqrt{n/k}/c_2$ we have that y is not in any of the balls around the different x_j with $|x_j| \geqslant R$. Therefore $K \cap E \subset R' D_n$ for $R' = R + r$, and we arrive at the conclusion of the theorem with, say, $C_1 = c_1^{-1/2}(25 C_0/c_2)$. □

Remark. The constant c_1 in the theorem above is determined by Chernoff, namely, we *must* have c_1 small enough to be able to use Chernoff for all of the vectors. In other words,

if r is defined as the smallest r' so that $N(K, r'D) \leqslant e^{c_1 k}$ (this is in fact called the $c_2 k$th entropy number of K) then we actually showed that $R = C'r\sqrt{n/k}$ is an upper bound for the radius of a random section of co-dimension k. We can even further improve since we need only cover $K \cap RD$, same as the standard way low M^* estimate is usually improved.

We now state and prove one slight generalization of this theorem which will be used in the next section. The following result, not yet derandomized, is for signed projections of bases.

Theorem 21. *Let K be a convex body in $E = \mathbb{R}^n$. Assume that this E is a subspace of a higher-dimensional space, that is, $E \subset R^N$, and let $\hat{e}_i = P_E e_i$ where $e_1, \ldots, e_N$ is some orthonormal basis in R^N. For a sequence of signs $\delta \in \{-1, 1\}^N$, define the vector $X(\delta) = \sum_{i=1}^N \delta_i \hat{e}_i$. If we choose $\xi_i = X(\varepsilon_i)$ where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k$ are random sign-vectors in $\{-1, 1\}^N$, then with probability greater than $1 - e^{-ck}$ one has for $F = \bigcap_{i=1}^k \operatorname{Ker}(\xi_i)$ that*

$$K \cap F \subset C_2 \frac{M^*}{(k/n)} D_n$$

(*where C_2 is universal*).

Proof. The proof goes along the same lines of the standard low M^* estimate. What one has only to notice is that for any $y \in E \subset R^N$

$$\mathbb{P}\big[\big|\langle y, \xi_j \rangle\big| \in [a, b]\big] = \mathbb{P}\left[\left|\left\langle y, \sum_{i=1}^N \varepsilon_i e_i \right\rangle\right| \in [a, b]\right]$$

and the latter is estimated with the use of the standard estimates (see, for example, [5]). This in fact goes in the direction of [28] where a larger class of distributions is considered. However, we only ever derandomize signed vectors, or more generally discrete probability densities, so there is no need to describe the general situation. □

The derandomized version works exactly in the same way, so we have

Theorem 22. *Let K be a convex body in $E = \mathbb{R}^n$. Assume that this E is a subspace of a higher-dimensional space, $E \subset R^N$, and let $\hat{e}_i = P_E e_i$ where $e_1, \ldots, e_N$ is some orthonormal basis in R^N. For a sequence of signs $\delta \in \{-1, 1\}^N$, define the vector $X(\delta) = \sum_{i=1}^N \varepsilon_i \hat{e}_i$. If we choose $\xi_i = X(\varepsilon_i)$ where the k vectors $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k$ are attained by taking a random walk from a uniform starting vertex on an expander with vertices $\{-1, 1\}^N$, with expansion parameter $n^{-\alpha}$ then with probability $1 - e^{-ck}$ one has for $F = \bigcap_{i=1}^k \operatorname{Ker}(\xi_i)$ that*

$$K \cap F \subset C_1 \frac{M^*(K)}{(k/n)} D_n$$

(*where C_1 and α are universal*).

Remark. Note that again we only use $N + k\alpha \log n$ random signs to generate the section.

9. The QS theorem

In this section we derandomize the QS-theorem of Milman [25], which states that for a body $K \subset \mathbb{R}^n$ there exist subspaces $F \subset E \subset \mathbb{R}^n$ both of dimension proportional to n, and when one takes first a projection onto E and then a section with F, one arrives at a body which is isomorphic to Euclidean up to a constant which depends only on the proportion of the co-dimensions of F and E to the dimension n of the full space. In fact, it was shown that this holds true with high probability on the choice of subspaces, provided the body K is in a "good position." That is, for any body K there is a linear transformation u so that the above holds with exponentially high probability for the body uK and random subspaces in the corresponding Grassmanians. In this section we prove the corresponding theorem for subspaces given as kernels of sign matrices, and derandomize it.

We start, however, with a weaker version, which is along the same lines of the original QS theorem's proof in [25]. Notice that in the formulation below the subspaces and quotients are given as kernels of sign-vectors. This is a new variant on the QS theorem.

Statement 23. *Let K be a symmetric convex body in $\mathbb{R}^n$ equipped with the standard Euclidean structure. Let $\xi_1, \xi_2, \ldots, \xi_k \in \{-1, 1\}^n$ and $\eta_1, \eta_2, \ldots, \eta_k \in \{-1, 1\}^n$ be chosen at random for $k = \varepsilon n$, $\varepsilon < 1/4$. Denote $E = \operatorname{Ker}(\xi_i, i = 1, \ldots, k)$ and $F = E \cap \operatorname{Ker}(\eta_i, i = 1, \ldots, k)$. Then with probability greater than $1 - e^{-ck}$ one has that*

$$d(P_E K \cap F, D_n \cap F) \leqslant C_0 M(K) M^*(K),$$

where C_0, c depend only on $\varepsilon = k/n$.

We remark that any body K has a position in which $M(K)M^*(K) \leqslant C \log(d(K, D_n))$ (see [32]).

Proof. To prove the statement, we simply use the low M^* result twice. So, the fact that with probability greater than $1 - e^{-ck}$ we have

$$K \cap E \subset C_1 \frac{M^*}{(k/n)} D_n$$

follows immediately from the low M^* theorem for signs (a special case of Theorem 21). This implies, by duality, that

$$\frac{k/n}{C_1 M^*} D_n \cap E \subset P_E(K^\circ).$$

Moreover, by standard considerations (see, e.g., [29]) we know that

$$M^*\big(P_E K^\circ\big) \leqslant \sqrt{\frac{n}{n-k}} M(K).$$

We now use Theorem 21 (because η_i come from the bigger space and are not in E) to get that again with high probability, $1-e^{-c'k}$, we have

$$P_E(K^\circ)\cap F \subset C_1\frac{\sqrt{n/(n-k)}M(K)}{(k/(n-k))} = C_1\frac{M(K)}{k/\sqrt{n(n-k)}}.$$

Joining the two inclusion relations we arrive at the desired result. □

Again derandomization is immediate and we have

Statement 24. *Let K be a symmetric convex body in $\mathbb{R}^n$ equipped with the standard Euclidean structure. Given $k<n/4$, there exists $\alpha(k/n)$ such that given $\alpha(k/n)\log n$ random sign vectors we can generate explicitly $\xi_1,\xi_2,\dots,\xi_k\in\{-1,1\}^n$ and $\eta_1,\eta_2,\dots,\eta_k\in\{-1,1\}^n$ (attained by taking two random walks from a uniform starting vertex each on an expander with vertices $\{-1,1\}^n$, with expansion parameter $n^{-\alpha}$) such that the following holds. Denote $E=\mathrm{Ker}(\xi_i, i=1,\dots,k)$ and $F=E\cap\mathrm{Ker}(\eta_i, i=1,\dots,k)$. Then with probability greater than $1-e^{-ck}$ one has that*

$$d(P_E K\cap F, D_n\cap F)\leqslant C_0 M(K)M^*(K),$$

where C_0, c, α depend only on k/n.

We turn now to the stronger version, which makes use of the M-position of a convex body. An M-position of a convex bodies has many equivalent definitions, and we state the properties which we use, and which can be used as a definition. We say that a convex body $K\subset\mathbb{R}^n$ is in M-position with constant C if $\mathrm{Vol}(K)=\mathrm{Vol}(D_n)$ and $\mathrm{Vol}(K+D)\leqslant C^n\,\mathrm{Vol}(D_n)$. This already implies that $\mathrm{Vol}(K\cap D)\geqslant C^{-n}\,\mathrm{Vol}(D)$; it also follows that the same is true, up to universal constants, for the dual body K°. It was shown by Milman [26] that every body has a linear image which is in M-position with a universal C. Below when we say that a body is in M-position we mean it is in M-position with constant C for a universal C. For details and proofs about M-position see [27,32].

The theorem which we first prove and then derandomize is the following.

Theorem 25. *Let K be a convex body in $\mathbb{R}^n$ equipped with the standard Euclidean structure. Assume further that K is in M-position. Take any orthogonal basis, and consider the discrete cube in this basis. Let $\xi_1,\xi_2,\dots,\xi_k\in\{-1,1\}^n$ and $\eta_1,\eta_2,\dots,\eta_k\in\{-1,1\}^n$ be chosen at random for $k=\varepsilon n$, $\varepsilon<1/4$. Denote $E=\mathrm{Ker}(\xi_i, i=1,\dots,k)$ and $F=E\cap\mathrm{Ker}(\eta_i, i=1,\dots,k)$. Then with probability greater than $1-e^{-ck}$ one has that*

$$d(P_E K\cap F, D_n\cap F)\leqslant C_0,$$

where c, C_0 depend only on k/n.

Proof. The fact that $K+D_n$ has finite volume ratio C implies immediately, using Theorem 19 in the random version, that for E defined as above with high probability

$(K+D_n)\cap E\subset C_1D_n$. Thus in particular $K\cap E\subset C_1D_n$ and so $P_EK^\circ\supset C_1^{-1}D\cap E$. The M-position also guarantees that P_EK° has finite volume ratio, since by Rogers–Shephard inequality one has

$$\operatorname{Vol}\bigl(P_EK^\circ\bigr)\operatorname{Vol}(K^\circ\cap E^\perp)\leqslant 2^n\operatorname{Vol}(K^\circ)$$

and by M-position

$$\operatorname{Vol}(K^\circ)\leqslant C^n\operatorname{Vol}\bigl(K^\circ\cap D_n\bigr)\leqslant C^n\operatorname{Vol}\bigl(P_E\bigl(K^\circ\cap D_n\bigr)\bigr)\operatorname{Vol}\bigl(K^\circ\cap D_n\cap E^\perp\bigr).$$

Joining this together with the trivial $\operatorname{Vol}(K^\circ\cap D_n\cap E^\perp)\leqslant\operatorname{Vol}(K^\circ\cap E^\perp)$ and $\operatorname{Vol}(P_E(K^\circ\cap D_n))\leqslant\operatorname{Vol}(P_ED_n)$, we get

$$\frac{\operatorname{Vol}(P_EK^\circ)}{\operatorname{Vol}(D_n\cap E)}\leqslant(2C)^n.$$

Recall that P_EK° already contains a ball of fixed radius. Therefore we can use the theorem for sections of finite volume ratio bodies again and have that the section of $P_E(K^\circ)$ by random F had bounded diameter which depends only on C and on k/n. In fact, we cannot simply use the random version of Theorem 19 because as before the section is by the kernel of a signed matrix in a larger space. We need a corresponding version of Theorem 19 for signed projections of basis. We give below as Statement 26 the derandomized version. Its proof, and the proof of the completely random version, are both almost identical to the proofs of Theorem 19 and of its completely random version, and are thus omitted. Thus, $P_EK^\circ\cap F$ is isomorphic to Euclidean up to constants depending only on k/n, and the proof is complete. □

Statement 26. *Let $E\subset\mathbb{R}^N$ with $\dim(E)=n$ and set $n/N=\delta$. Let e_i denote some orthonormal basis in $\mathbb{R}^N$ and $\hat{e}_i=P_Ee_i$, for $i=1,\dots,N$. For any body $K\subset E$ with finite volume ratio A, setting it in the position where the maximal volume ellipsoid of K is $D_N\cap E$, for any $\varepsilon>0$ we can use $C_0(A,\delta)\log n$ random sign vectors in $\{-1,1\}^N$ (i.e., $C_0(A,\delta)N\log n$ random bits) to create explicitly $k=\varepsilon n$ vectors $\xi_1,\xi_2,\dots,\xi_k$ of the form $\xi_j=\sum_{i=1}^N\pm\hat{e}_i$, satisfying that for any $y\in\mathbb{R}^n$ with $\langle y,\xi_i\rangle=0$ for every i, we have*

$$c_1(\varepsilon,A,\delta)|y|\leqslant\|y\|_K\leqslant|y|.$$

The derandomized version of Theorem 25, now using Theorem 19 itself which is derandomized, and the statement above which is a slight generalization of it we arrive at the following theorem.

Theorem 27. *Let K be a convex body in $\mathbb{R}^n$ equipped with the standard Euclidean structure. Assume that K is in M-position. We can use $C(k/n)\log n$ random sign-vectors to generate, explicitly, $2k$ vectors $\xi_1,\xi_2,\dots,\xi_k\in\{-1,1\}^n$ and $\eta_1,\eta_2,\dots,\eta_k\in\{-1,1\}^n$,*

for $k = \varepsilon n$, $\varepsilon < 1/4$, so that the following holds. Denote $E = \mathrm{Ker}(\xi_i, i = 1, \ldots, k)$ and $F = E \cap \mathrm{Ker}(\eta_i, i = 1, \ldots, k)$. Then with probability greater than $1 - e^{-ck}$ one has that

$$d(P_E K \cap F, D_n \cap F) \leqslant C_0,$$

where $c, C_0, C(\varepsilon)$ depend only on $\varepsilon = k/n$.

Note added in proof

It has come to our attention that, in the framework of Computer Science, there are few results regarding large sections of ℓ_1 determined by small number of bits (see [16] for $n(\log n)^2$ bits and references therein). However, it is emphasized in [16] that their presentations are nonconstructive which corresponds, in our understanding, to "existence" results.

Acknowledgments

We thank Avi Wigderson for introducing to us the subject of derandomization using random walks on expander graphs, and for fruitful conversations. We also thank Noga Alon for fruitful and inspiring discussions.

References

[1] M. Ajtai, J. Komlós, E. Szemerédi, Deterministic simulation of logspace, in: Proc. 19th ACM Sympos. on Theory of Computing, ACM, 1987, pp. 132–140.
[2] N. Alon, U. Feige, A. Wigderson, D. Zuckerman, Derandomized graph products, Comput. Complexity (1995) 60–75.
[3] N. Alon, V. Milman, λ_1, isoperimetric inequalities for graphs and superconcentrators, J. Combin. Theory Ser. B 38 (1985) 73–88.
[4] N. Alon, J. Spencer, The Probabilistic Method, second ed., with an appendix on the life and work of Paul Erdős, Wiley–Intersci. Ser. Discrete Math. Optim., Wiley, New York, 2000.
[5] S. Artstein, The change in the diameter of a convex body under a random sign-projection, in: Geometric Aspects of Functional Analysis, in: Lecture Notes in Math., vol. 1850, Springer, Berlin, 2004, pp. 31–39.
[6] S. Artstein-Avidan, O. Friedland, V. Milman, Some geometric applications of Chernoff-type estimates, in: Geometric Aspects of Functional Analysis, in: Lecture Notes in Math., Springer, Berlin, 2006, in press.
[7] S. Artstein-Avidan, O. Friedland, V. Milman, Geometric applications of Chernoff-type estimates and a ZigZag approximation for balls, Proc. Amer. Math. Soc., in press.
[8] A. Ben-Tal, A. Nemirovski, On polyhedral approximations of the second-order cone, Math. Oper. Res. 26 (2) (2001) 193–205.
[9] J. Bourgain, J. Lindenstrauss, V. Milman, Minkowski sums and symmetrizations, in: Geometric Aspects of Functional Analysis (1986/1987), in: Lecture Notes in Math., vol. 1317, Springer, Berlin, 1988, pp. 44–66.
[10] G. Cheang, A. Barron, A better approximation for balls, J. Approx. Theory 104 (2) (2000) 183–302.
[11] A. Cohen, A. Wigderson, Dispersers, deterministic amplification, and weak random sources, in: Proc. 30th IEEE Sympos. on Foundations of Computer Science, IEEE, 1989, pp. 14–19.
[12] T. Figiel, J. Lindenstrauss, V. Milman, The dimension of almost spherical sections of convex bodies, Acta Math. 139 (1–2) (1977) 53–94.

[13] D. Gillman, A Chernoff bound for random walks on expander graphs, SIAM J. Comput. 27 (4) (1998) 1203–1220.
[14] Y. Gordon, On Milman's inequality and random subspaces which escape through a mesh in R^n, in: Geometric Aspects of Functional Analysis (1986/1987), in: Lecture Notes in Math., vol. 1317, Springer, Berlin, 1988, pp. 84–106.
[15] R. Impagliazzo, D. Zuckerman, How to reduce random bits, in: Proc. 30th IEEE Sympos. on Foundations of Computer Science, IEEE Press, 1989, pp. 248–253.
[16] P. Indyk, Stable distributions, pseudorandom generators, embeddings and data stream computation, in: 41st Symposium on Foundations of Computer Science, 2000, pp. 189–197.
[17] H. Iwaniec, E. Kowalski, Analytic Number Theory, Amer. Math. Soc. Colloq. Publ., vol. 53, Amer. Math. Soc., Providence, RI, 2004.
[18] B. Kashin, Section of some finite-dimensional sets and classes of smooth functions, Izv. Akad. Nauk SSSR 41 (1977) 334–351 (in Russian).
[19] A. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, R. Vershynin, Random Euclidean embeddings in spaces of bounded volume ratio, C. R. Acad. Sci. Paris Sér. I Math. 339 (1) (2004) 33–38.
[20] A. Lubotzky, Discrete Groups, Expanding Graphs and Invariant Measures, with an appendix by J.D. Rogawski, Progr. Math., vol. 125, Birkhäuser, Basel, 1994.
[21] A. Lubotzky, A. Phillips, P. Sarnak, Ramanujan Graphs, Combinatorica 8 (1988) 261–277.
[22] G.A. Margulis, Explicit constructions of expanders, Problemy Peredachi Informatsii 9 (4) (1973) 71–80 (in Russian).
[23] G.A. Margulis, Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and superconcentrators, Problems Inform. Transmission 24 (1988) 39–46.
[24] V.D. Milman, Random subspaces of proportional dimension of finite-dimensional normed spaces: approach through the isoperimetric inequality, in: Proceedings, Missouri, in: Lecture Notes in Math., vol. 1166, Springer, Berlin, 1984, pp. 106–115.
[25] V.D. Milman, Almost Euclidean quotient spaces of subspaces of finite-dimensional normed spaces, Proc. Amer. Math. Soc. 94 (1985) 445–449.
[26] V.D. Milman, Inegalité de Brunn–Minkowski inverse et applications à la théorie locale des espaces normés, C. R. Acad. Sci. Paris Sér. I Math. 302 (1986) 25–28.
[27] V.D. Milman, Isomorphic symmetrizations and geometric inequalities, in: Geometric Aspects of Functional Analysis (1986/1987), in: Lecture Notes in Math., vol. 1317, Springer, Berlin, 1988, pp. 107–131.
[28] V.D. Milman, A. Pajor, Regularization of star bodies by random hyperplane cut off, Studia Math. 159 (2) (2003) 247–261.
[29] V.D. Milman, G. Schechtman, Asymptotic Theory of Finite-Dimensional Normed Spaces, with an appendix by M. Gromov, Lecture Notes in Math., vol. 1200, Springer, Berlin, 1986.
[30] V.D. Milman, G. Schechtman, Global versus local asymptotic theories of finite-dimensional normed spaces, Duke Math. J. 90 (1) (1997) 73–93.
[31] A. Pajor, N. Tomczak-Jaegermann, Subspaces of small codimension of finite-dimensional Banach spaces, Proc. Amer. Math. Soc. 97 (4) (1986) 637–642.
[32] G. Pisier, The Volume of Convex Bodies and Banach Space Geometry, Cambridge Tracts in Math., vol. 94, Cambridge Univ. Press, Cambridge, 1989.
[33] P. Sarnak, Some Applications of Modular Forms, Cambridge Tracts in Math., vol. 99, Cambridge Univ. Press, Cambridge, 1990.
[34] G. Schechtman, Random embeddings of Euclidean spaces in sequence spaces, Israel J. Math. 40 (2) (1981) 187–192.
[35] G. Schechtman, Special orthogonal splittings of L_1^{2k}, Israel J. Math. 139 (2004) 337–347.
[36] C. Schütt, Entropy numbers of diagonal operators between symmetric Banach spaces, J. Approx. Theory 40 (2) (1984) 121–128.