Polynomial bounds for large Bernoulli sections of ℓ_1^N

S. Artstein-Avidan[1,2], O. Friedland[3], V. Milman[1,3], S. Sodin[3]

1 Introduction

This paper consists of two distinct parts. The first one presents the "local" version of the result of Bai and Yin from [2]. This result gives an estimate from below for the probability that the smallest singular value of a random sign matrix is outside some interval. In particular, it gives a lower bound for the probability that an "almost square" matrix, that is, a $(1-\delta)n \times n$ matrix, has smallest singular value above $\approx \delta$. This is a "finite dimensional" version of the results of Bai and Yin [2], and in this "local" version it is much more useful for applications in Asymptotic Geometric Analysis problems, where quantitative estimates of deviations are needed. This is presented in Section 2. A more extensive presentation of this result will be given by the IVth named author in [14].

The second part of this paper consists of precisely such an application,

[1]partially supported by BSF grant 2002-006
[2]supported by the National Science Foundation under agreement No. DMS-0111298
[3]supported in part by the Israel Science Academy

where the method of [1] and some other recent developments are joined with the above, to improve results from [8] and from [1] regarding the distance from euclidean space of almost full dimensional sections of the space ℓ_1^N realized as images of sign matrices. For $N = (1+\delta)n$ we receive estimates on the isomorphism constant which are much better than were previously known, and in particular are *polynomial* in δ.

Acknowledgments The authors thank Prof. S. Szarek for bringing the paper [2] to their attention and suggesting that the methods used there yield an estimate on the probability of deviation. The IVth named author thanks Prof. E. Gluskin for useful discussions of the combinatorial estimates in Section 2.3. The authors thank the referee for his careful reading and useful comments.

2 The rate of convergence in the result of Bai and Yin

In this section we present a lower bound on the least singular value of a Bernoulli random matrix, in the spirit of Bai and Yin [2].

2.1 Introduction and main statement

Let X be a $p \times n$ matrix of random signs: X_{ik} are independent for $1 \le i \le p$ and $1 \le k \le n$,

$$\mathbb{P}\{X_{ik} = 1\} = \mathbb{P}\{X_{ik} = -1\} = 1/2\,. \tag{1}$$

We study the spectrum Λ_S of the covariance matrix

$$S = n^{-1} X X^{\mathrm{T}} . \tag{2}$$

Let $\mu_S = p^{-1} \sum_{\lambda \in \Lambda(S)} \delta_\lambda$ be the empirical eigenvalue distribution of S. Marchenko and Pastur [9] proved that

$$d\mu_S \xrightarrow{\text{a.s.}} f_{\mathrm{MP}} dx \quad \text{as} \quad n \to \infty ,$$

where the limit density equals

$$f_{\mathrm{MP}}(x) = \begin{cases} \frac{1}{2\pi y x} \sqrt{(x-a)(b-x)}, & a \le x \le b \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

with the notation

$$y = p/n < 1 \text{ (fixed)}, \quad a = (1-\sqrt{y})^2, \quad b = (1+\sqrt{y})^2 . \tag{4}$$

It is natural to ask whether the eigenvalues of S can lie far from the support $[a, b]$ of this distribution. Bai and Yin [2] answered negatively, proving (for a more general random matrix model) that with probability 1

$$\lambda_{\min}(S) \to a, \quad \lambda_{\max}(S) \to b \quad \text{as } n \to \infty. \tag{BY}$$

In the spirit of local theory we strive for a quantitative form of this result.

Theorem 1. *There exists a universal constant $C > 0$ such that the following holds. Let X be a $p \times n$ matrix of random signs as defined by (1); define S as in (2) and y, a and b as in (4); assume that*

$$\frac{C \log^2 n}{\sqrt{y} \sqrt[3]{n}} \le \epsilon \le 1 . \tag{5}$$

Then the probability that S has eigenvalues outside $[a-\epsilon, b+\epsilon]$ is less than

$$\exp\left(-C^{-1} y^{1/6} n^{1/6} \epsilon^{1/2}\right) = \exp\left(-C^{-1} p^{1/6} \epsilon^{1/2}\right) .$$

For y close to 1 the theorem yields the following lower bound on the least eigenvalue of S:

Theorem 2. *There exists a universal constant $C > 0$ such that if, in the notation of Theorem 1, $y = 1 - \delta$ with $1/2 > \delta > Cn^{-1/6} \log n$, then*

$$\mathbb{P}\left\{\lambda_{\min}(S) \leq \delta^2/8\right\} \leq \exp\left(-C^{-1} n^{1/6} \delta\right) .$$

Recently, Litvak, Pajor, Rudelson and Tomczak-Jaegermann [7] proved (in a more general setting) that if $y = 1 - \delta$ with $1 > \delta \geq c_1 / \ln c_2 n$ in the notation of Theorem 1, then

$$\mathbb{P}\left\{\lambda_{\min}(S) \leq A a^{1/\delta}\right\} \leq \exp(-Cn) , \qquad \text{(LPRT)}$$

where $A > 0$, $1 > a > 0$, $C, c_1, c_2 > 0$ are universal constants.

Note that the bound on the probability decays exponentially; this is rather important in geometric applications. We do not know whether the left-hand side in Theorem 2 is in fact as small as $\exp\left(-n\delta^C / C\right)$ for some $C > 0$.

Let us show that Theorem 1 implies Theorem 2.

Proof of Theorem 2. The Taylor expansion yields $\sqrt{y} \approx 1 - \delta/2$ and hence

$$(1 - \sqrt{y})^2 - \epsilon \approx \delta^2/4 - \epsilon .$$

Now take $\epsilon \approx \delta^2/8$ and use Theorem 1. We obtain:

$$\mathbb{P}\left\{\lambda_{\min}(S) < \delta^2/8\right\} \leq \exp\left(-\frac{\sqrt[6]{y}}{\sqrt{8}C} n^{1/6} \delta\right) \leq \exp\left(-\frac{n^{1/6}\delta}{2^{5/3}C}\right) .$$

□

The main idea behind the proof of Theorem 1 makes use of the following construction, due to Bai and Yin [2]. We define a sequence of matrices $T(l)$, $l = 0, 1, 2, \cdots$, that are certain polynomials of the matrix $T = S - \mathbb{I}$: $T(l) = p_l(T)$. If $\mu_1, \cdots, \mu_p$ are the eigenvalues of T, then $p_l(\mu_1), \cdots, p_l(\mu_p)$ are the eigenvalues of $T(l)$.

The polynomials p_l can be expressed via the Chebyshev polynomials of the second kind. If $\mu \notin [a-1,\, b-1]$, the sequence $p_l(\mu)$ tends to infinity exponentially fast. We define p_l and prove these observations in Section 2.2.

On the other hand, the expression $\mathbb{E}\,\mathrm{Tr}\,T(l)$ allows a graph-theoretical interpretation showing that it can not grow too fast. We prove such a bound in Section 2.3, using a modification of the combinatorial argument due to Bai and Yin.

In Section 2.4 we combine these facts and obtain a bound on $a - \lambda_{\min}(S)$, $\lambda_{\max}(S) - b$ that concludes the proof of Theorem 1.

2.2 Construction and basic properties of $T(l)$

Denote $y_1 = \frac{p-2}{n}$, $y_2 = \frac{(p-1)(n-1)}{n^2}$; $y \geq y_2 \geq y_1 = y - \frac{2}{n}$.

Define a sequence of matrices $T(l) = (T_{ij}(l))_{ij}$,

$$\begin{cases} T(0) = \mathbb{I}, \quad T(1) = T = n^{-1} X X^{\mathrm{T}} - \mathbb{I}, \\ T(l+1) = (T - y_1 \mathbb{I}) \cdot T(l) - y_2 \cdot T(l-1)\,. \end{cases} \tag{6}$$

We have: $T(l) = p_l(T)$, where

$$\begin{cases} p_0(\mu) = 1, \quad p_1(\mu) = \mu, \\ p_{l+1}(\mu) = (\mu - y_1) \cdot p_l(\mu) - y_2 \cdot p_{l-1}(\mu)\,. \end{cases}$$

Recall the definition

$$U_l(\cos\theta) = \frac{\sin((l+1)\,\theta)}{\sin\theta} \tag{Cheb1}$$

of the Chebyshev polynomials of the second kind. Here, both the right-hand side and the left-hand side are polynomials; hence the equality makes sense for any $\theta \in \mathbb{C}$.

Equivalently, U_l can be defined by the recurrence relation

$$\begin{cases} U_0(x) = 1, \quad U_1(x) = 2x, \\ U_{l+1}(x) = 2xU_l(x) - U_{l-1}(x)\,. \end{cases} \tag{Cheb2}$$

The latter definition readily yields the formula

$$p_l(\mu) = {y_2}^{l/2} U_l\left(\frac{\mu - y_1}{2\sqrt{y_2}}\right) + y_1 {y_2}^{(l-1)/2} U_{l-1}\left(\frac{\mu - y_1}{2\sqrt{y_2}}\right)\,. \tag{7}$$

Remark. If we replace y_1 and y_2 with y in (7), the sequence becomes orthogonal with respect to the Marchenko–Pastur measure (3). Kusalik, Mingo and Speicher [6] used a different form of this sequence to study the spectral properties of random matrices with complex Gaussian entries, and called it the sequence of *shifted Chebyshev polynomials of the second kind*.

Now we use (Cheb1) to estimate the polynomials p_l.

Lemma 3. *There exists a universal constant $C > 0$ such that the following properties hold for any even $l \geq 2$, $0 \leq \epsilon \leq 1$:*

1. *For any $\mu \in \mathbb{R}$,*

$$p_l(\mu) \geq -2ly^{l/2}\,. \tag{8}$$

2. If

$$|\mu - y_1| \geq 2\sqrt{y_2} + \epsilon \,,$$

then

$$p_l(\mu) \geq {y_2}^{l/2} \exp\left(C^{-1} l \epsilon^{1/2} {y_2}^{-1/4}\right) \,. \tag{9}$$

Proof.

1. If x lies outside the interval $[\cos\frac{l\pi}{l+1}, \cos\frac{\pi}{l+1}]$, then $U_l(x) > 0$. Otherwise, $x = \cos\theta$ for some $\frac{\pi}{l+1} \leq \theta \leq \frac{l\pi}{l+1}$; therefore

$$U_l(x) \geq -\sin^{-1}\frac{\pi}{l+1} \geq -\frac{l+1}{2}\,.$$

Hence

$$p_l(\mu) \geq -\left({y_2}^{l/2} + y_1 {y_2}^{(l-1)/2}\right)\frac{l+1}{2} \geq -2y^{l/2} l\,.$$

2. If $|x| \geq 1 + \epsilon$, $x = \cos i\theta$ for some $\theta \geq C^{-1}\epsilon^{1/2}$; hence

$$U_l(x) = \frac{\sin((l+1)\, i\theta)}{\sin i\theta} \geq e^{l\theta/2} \geq e^{C_1^{-1} l \epsilon^{1/2}}\,.$$

Therefore

$$p_l(\mu) \geq {y_2}^{l/2} \exp\left(\frac{l}{C}\frac{\sqrt{\epsilon}}{\sqrt[4]{y}}\right)\,.$$

□

Next we apply (8) and (9) to the eigenvalues of T.

Lemma 4. *There exists a universal constant $C > 0$ such that if $n \geq l \geq 2$ is even,*

$$C \max\left(\frac{1}{\sqrt{y} n}\,, \frac{\sqrt{y}\log^2 n}{l^2}\right) \leq \epsilon \leq 1 \qquad \textit{and} \tag{10}$$

$$\max\left\{|\mu - y| \,\Big|\, \mu \textit{ is an eigenvalue of } T\right\} \geq 2\sqrt{y} + \epsilon\,,$$

then

$$\operatorname{Tr} T(l) \geq y^{l/2} \exp\left(C^{-1} l \epsilon^{1/2} y^{-1/4}\right)\,. \tag{11}$$

Proof. Let $\mu_1, \cdots, \mu_p$ be the eigenvalues of T; suppose $|\mu_1 - y| \geq 2\sqrt{y} + \epsilon$.

Then by (10)

$$\begin{aligned} |\mu_1 - y_1| &\geq 2\sqrt{y_2} + \epsilon - 2\left[\sqrt{y} - \sqrt{y_2}\right] - [y - y_1] \\ &\geq 2\sqrt{y_2} + \epsilon - \frac{4}{n\sqrt{y}} - \frac{2}{n} \\ &\geq 2\sqrt{y_2} + \epsilon - \frac{6}{n\sqrt{y}} \geq 2\sqrt{y_2} + C_1\epsilon\,. \end{aligned}$$

Write the bound (9) with $\mu = \mu_1$ and the bound (8) with $\mu = \mu_2, \cdots, \mu_p$; add the inequalities and use (10) once again:

$$\begin{aligned} \operatorname{Tr} T(l) &\geq {y_2}^{l/2} \exp\left(C^{-1} l \frac{\sqrt{C_1 \epsilon}}{\sqrt[4]{y_2}}\right) - 2lp\, y^{l/2} \\ &\geq C_2^{-1} y^{l/2} \exp\left(C_2^{-1} l \frac{\sqrt{\epsilon}}{\sqrt[4]{y}}\right) - 2n^2 y^{l/2} \\ &\geq y^{l/2} \exp\left(C_3^{-1} l \epsilon^{1/2} y^{-1/4}\right)\,. \end{aligned}$$

□

2.3 Combinatorial description of $T(l)$

Now we give a combinatorial description of $\mathbb{E} \operatorname{Tr} T(l)$.

Lemma 5. *The following equality holds:*

$$T_{ij}(l) = \frac{1}{n^l} \sum\nolimits^{*} X_{iv_1} X_{u_1 v_1} X_{u_1 v_2} X_{u_2 v_2} \cdots X_{u_{l-1} v_l} X_{j v_l}\,, \tag{12}$$

where the sum $\sum^*$ *is over all* $u_1, \ldots, u_{l-1}$ *and* $v_1, \ldots, v_l$ *satisfying* $1 \le u_r \le p$ *for* $1 \le r \le l-1$ *and* $1 \le v_s \le n$ *for* $1 \le s \le l$*, and such that, in addition,*

$$\begin{cases} i \ne u_1 \ne u_2 \ne u_3 \ne \cdots \ne u_{l-1} \ne j \\ v_1 \ne v_2 \ne v_3 \ne \cdots \ne v_l \end{cases} .$$

(Notice that there is no requirement $u_1 \ne u_3$, for example.)

Proof. Denote by $T'_{ij}(l)$ the right-hand side of (12); denote $T'(l) = (T'_{ij}(l))$. Then $T'(0) = \mathbb{I} = T(0)$, $T'(1) = T = T(1)$.

Further, $(T \cdot T'(l-1))_{ij}$ is a sum of the same form as (12), but without the condition $v_1 \ne v_2$. The three cases (i) $v_1 \ne v_2$, (ii) $v_1 = v_2$ and $i \ne u_2$, (iii) $v_1 = v_2$ and $i = u_2$ yield the terms

$$T'(l)\,, \quad y_1\, T'(l-1)\,, \quad y_2\, T'(l-2)\,,$$

respectively. Therefore $T'(l)$ satisfy the same recurrence relation (6) as $T(l)$; this concludes the proof. ☐

The random variables X_{uv} are independent; therefore the expectation of a term in (12) vanishes unless every X_{uv} appears an even number of times in the product. In the latter case, the expectation equals 1 (note that 0 is even).

Corollary 6. *The expectation* $n^l\, \mathbb{E}\, \mathrm{Tr}\, T(l)$ *equals the number of configurations*

$$1 \le i, u_1, u_2, u_3, \cdots, u_{l-1} \le p, \quad 1 \le v_1, v_2, \cdots, v_l \le n\,,$$

such that

$$\begin{cases} i \neq u_1 \neq u_2 \neq u_3 \neq \cdots \neq u_{l-1} \neq i \\ v_1 \neq v_2 \neq v_3 \neq \cdots \neq v_l \end{cases},$$

and every pair uv appears an even number of times in the sequence

$$iv_1, u_1v_1, u_1v_2, u_2v_2, \cdots, u_{l-1}v_l, iv_l \, .$$

The following graph-theoretical interpretation will be of use. Every configuration of i, u_r and v_s which is permitted in Corollary 6 corresponds to a closed path W in the bipartite graph $K_{p,n}$ such that

(W1) the path W passes through every edge an even number of times;

(W2) W never passes through an edge 2 times consequently (i.e. the pattern $w \to w' \to w$ is not allowed).

(Moreover, every path begins on the left side of the graph, but we ignore this in our estimates.)

Let W be a closed path on an arbitrary graph G so that (W1) and (W2) hold. Consider W as a set of triples (w_1, w_2, r), where $1 \leq r \leq 2l$, meaning that the rth edge on W goes from w_1 to w_2.

Divide the edges into 3 classes:

$$T_1 = \big\{(w_1, w_2, r) \in W \,\big|\, \forall r' < r, (w_1', w_2', r') \in W \Rightarrow w_1' \neq w_2 \wedge w_2' \neq w_2\big\} \, ,$$

$$T_2 = \big\{(w_1, w_2, r) \in W \,\big|\, \exists r' < r : (w_1, w_2, r') \in T_1 \vee (w_2, w_1, r') \in T_1 \, , \; \forall r' < r'' < r : (w_1, w_2, r'') \notin W \wedge (w_2, w_1, r'') \notin W\big\} \, ,$$

$$T_3 = W \setminus (T_1 \cup T_2) \, .$$

(*Semiformal verbal description:* The edges of T_1 are the first edges to visit their endpoints; that is, T_1 is the DFS tree of W. Every edge in T_1 appears at least once again on W; we denote by T_2 the set of second appearances of the T_1 edges. All the other edges form the set T_3.)

Let us call a sequence of vertices $f = w_1 w_2 \cdots w_k$ $(k > 1)$ a *protofragment* of W if the following 3 conditions hold: *(i)* for some r

$$(w_1, w_2, r), (w_2, w_3, r+1), \cdots, (w_{k-1}, w_k, r+k-2) \in T_1 ,$$

(ii) for some r'

$$\begin{cases} \text{either} \quad (w_1, w_2, r'),\ (w_2, w_3, r'+1),\ \cdots,\ (w_{k-1}, w_k, r'+k-2) \in T_2 \\ \text{or} \qquad (w_k, w_{k-1}, r'),\ \cdots,\ (w_3, w_2, r'+k-3),\ (w_2, w_1, r'+k-2) \in T_2 , \end{cases}$$

and *(iii)* f is maximal with respect to the 2 conditions *(i)–(ii)*.

If $f = w_1 w_2 \cdots w_k$ is a protofragment, $w_1 \neq i$, we call its suffix $\bar{f} = w_2 \cdots w_k$ a *fragment* of length $k-1$. If $w_1 = i$, we call f a *fragment* of length k. The vertices of W are thereby divided into F fragments.

The following combinatorial bound will be crucial ($\sharp$ denotes cardinality):

Lemma 7. $F \leq 2\sharp T_3 + 1$.

Proof. Let f be a protofragment that starts with $w \neq i$; consider 2 cases. If f is passed in the same direction in T_1 and T_2, the edge adjacent to w in one of the 2 passages is in T_3.

Otherwise, the last edge before the second appearance of f is in T_3.

Let$e$ be the T_3 edge in either case. The map $f \mapsto e$ is at most 2–1; hence $F - 1 \leq 2\sharp T_3$.

□

Lemma 8. *The number of different fragments of length k in $K_{p,n}$ is bounded by $2y^{-1/2}(pn)^{k/2}$.*

Proof. First decide to which side of the graph does the first vertex belong. Then choose all the vertices. □

Now we can bound the number of paths satisfying (W1)–(W2) on $K_{p,n}$. Let V be the number of (distinct) vertices on W.

First, choose the lengths of the fragments. This can be done in $\binom{V}{F-1} \le V^F/F!$ ways. Next, choose the fragments themselves; by Lemma 8 this can be done in at most $(y/4)^{-F/2}(pn)^{V/2}$ ways.

We can change the directions of the fragments in T_2, in 2^F ways. Now that the fragments are ready, glue them onto the path; this can be done in $(2l-2V+1)^{2F}$ ways (just pick a place for every fragment).

Now there are $2l-2V$ vertices left. Every one of them coincides with one of the V vertices that we already have. Once we choose one of the V^{2l-2V} arrangements, our path is ready.

Multiplying all these numbers, we see that the number $\mathcal{P}$ of paths is bounded by

$$\begin{aligned}\mathcal{P} &\le \sum_{V=1}^{l}\sum_{F=1}^{l}\frac{V^F}{F!}(y/4)^{-F/2}(pn)^{V/2}2^F(2l-2V+1)^{2F}V^{2l-2V}\\ &\le \sum_{V=1}^{l}\sum_{F=1}^{l}(pn)^{V/2}\left(CVy^{-1/2}\right)^F V^{2l-2V}\times\left(\frac{2l-2V+1}{F}\right)^F .\end{aligned}$$

Now, $(x/F)^F \le e^x$; $F \le 2\sharp T_3 + 1 = 4l - 4V + 5$. Therefore

$$
\begin{aligned}
\mathcal{P} &\le \sum_{V=1}^{l}\sum_{F=1}^{l}(pn)^{V/2}(C'Vy^{-1/2})^{4l-4V+5}V^{2l-2V} \\
&\le \sum_{V=1}^{l} l\,(pn)^{V/2}(C'V^{3/2}y^{-1/2})^{4l-4V+5} \\
&\le l^9 y^{-5/2}(pn)^{l/2}\sum_{V=1}^{l}(pn)^{(V-l)/2}(C'V^{3/2}y^{-1/2})^{4l-4V}\,.
\end{aligned}
$$

Now, if $(C'l^{3/2}y^{-1/2})^8 < pn$, every term in the sum is no greater than 1. Therefore if

$$l < C''^{-1}y^{1/3}(pn)^{1/12} = C''^{-1}y^{5/12}n^{1/6}\,,$$

then

$$\mathcal{P} \le l^{10}y^{-5/2}(pn)^{l/2}\,;$$

finally (in one line):

$$\mathbb{E}\,\mathrm{Tr}\,T(l) \le l^{10}y^{(l-5)/2} \quad \text{if } l \le C''^{-1}y^{5/12}n^{1/6}. \tag{13}$$

2.4 Conclusion of the proof

Proof of Theorem 1. Let $l = 2\lfloor (2C'')^{-1}y^{5/12}n^{1/6}\rfloor$ in (13).

Then by (5)

$$\epsilon \ge \frac{C\log^2 n}{\sqrt{y}\sqrt[3]{n}} \ge \frac{C}{\sqrt{y}n}$$

and

$$\epsilon \ge \frac{C\log^2 n}{\sqrt{y}\sqrt[3]{n}} \ge \frac{C\sqrt{y}\log^2 n}{l^2}\frac{l^2}{y\sqrt[3]{n}} \ge \frac{C_1\sqrt{y}\log^2 n}{l^2}\,;$$

therefore (10) holds.

By Lemma 4, Chebyshev's inequality, the estimate (13) and the condition (5) that we imposed on ϵ,

$$
\begin{aligned}
&\mathbb{P}\{S \text{ has eigenvalues outside } [a-\epsilon,\, b+\epsilon]\} \\
&\qquad = \mathbb{P}\{T \text{ has eigenvalues outside } [a-1-\epsilon,\, b-1+\epsilon]\} \\
&\qquad \le \mathbb{P}\left\{\operatorname{Tr} T(l) \ge y^{l/2} \exp\left(C^{-1} l \epsilon^{1/2} y^{-1/4}\right)\right\} \\
&\qquad \le \frac{\mathbb{E} \operatorname{Tr} T(l)}{y^{l/2} \exp\left(C^{-1} l \epsilon^{1/2} y^{-1/4}\right)} \le C y^{5/3} n^{5/3} \exp\left(-C^{-1} y^{1/6} n^{1/6} \epsilon^{1/2}\right) \\
&\qquad \le \exp\left(-{C_1}^{-1} n^{1/6} y^{1/6} \epsilon^{1/2}\right) = \exp\left(-{C_1}^{-1} p^{1/6} \epsilon^{1/2}\right) .
\end{aligned}
$$

We are done.

□

3 Application to large sections of ℓ_1^N

It is well known that $\ell_1^{(1+\delta)n}$ has isomorphic euclidean sections of dimension n (see [5]), with constant of isomorphism independent of the dimension n and depending only on δ. When the section is taken to be the image of a matrix with i.i.d. gaussian entries (which is the same as taking a random subspace in the Grassmanian $G_{N,n}$ with respect to the normalized Haar measure), the dependence is polynomial in δ, with high probability on the choice of the entries. This was discovered first in the results of [3].

The image of a matrix of signs is simply the span of some set of vertices of the unit cube, and thus has more structure, and is sometimes more useful in implementations. Schechtman showed in [11] that the image of a matrix whose rows are $N = Cn$ sign-vectors in $\mathbb{R}^n$, where C is a universal constant, also realizes, with high probability on the choice of signs, an isomorphic to euclidean section of ℓ_1^N. The question then remained whether the constant C can be reduced to be close to 1. This was resolved by Johnson and Schechtman, and follows from their paper [4]. However, they showed the *existence* of such a sign matrix, and not that it is satisfied for a matrix whose rows are $N = (1+\delta)n$ *random* sign-vectors. In a paper by Litvak, Pajor, Rudelson, Tomczak-Jaegermann and Vershynin [8] this was demonstrated. However, the dependence of the constant of isomorphism on δ in their result is exponentially bad, and they get $c(\delta) = c^{1/\delta}$. In this paper we get a better dependence, polynomial in δ, however the probability that we get is slightly smaller than the probability in [8], with $n^{1/6}$ in the exponent instead of n.

We remark that results of this type can be viewed also in a different way, as a realization of Khinchine inequality with few vectors. The classical Khinchine inequality states that (for best constants as below see [15])

$$\frac{1}{\sqrt{2}}(\sum_{i=1}^n x_i^2)^{1/2} \le Ave_{\varepsilon_1,\dots,\varepsilon_n\ =\pm 1}|\sum_{i=1}^n \varepsilon_i x_i| \le (\sum_{i=1}^n x_i^2)^{1/2}.$$

Instead of averaging over *all* sign-vectors we may average over only $n(1+\delta)$ of them (chosen randomly, and good for *all* x), and get the same inequality only with a worse constant instead of $\sqrt{2}$. The constant is universal for fixed δ, and the way it behaves when $\delta \to 0$ is the subject of this paper, reformulated.

In this section we show that for a random $N \times n$ sign matrix, where $N = n(1+\delta)$, we have with high probability that the section of ℓ_1^N given by its image is isomorphic to the euclidean ball with polynomial dependence of the constants of isomorphism on δ. The developments which allowed this advancement include the methods of Schechtman to get L_1 splitting as in [12], the quantitative version of the result of Bai and Yin [2] given in Theorem 2 of the previous section, and the use of Chernoff bounds for geometric purposes much like is done in [1]. We prove

Theorem 9. *There exist universal constants* δ_0, c', c'', *and* c_0 *such that the following holds. Let* $c''n^{-1/6}\log n < \delta < \delta_0$, *and denote* $N = (1+\delta)n$. *Then with probability greater than* $1 - e^{-c'\delta n^{1/6}}$, *for* $(1+\delta)n$ *random sign-vectors* $\varepsilon_j \in \{-1,1\}^n$, $j = 1, \ldots, n+\delta n$, *one has for every* $x \in \mathbb{R}^n$

$$c(\delta)|x| \le \frac{1}{N}\sum_{j=1}^{N} |\langle x, \varepsilon_j\rangle|, \tag{14}$$

where $c(\delta) = c_0\delta^{5/2}/\log(1/\delta)$.

In fact it is easy to see that once we know Theorem 9 the above remains true for any $\delta > 0$, and the restriction $\delta < \delta_0$ is artificial. Also, an upper bound in (14) is known and standard, similar to Lemma 11.

Notation: We pick the $N = n+\delta n$ random sign vectors ε_j, normalize them to be unit vectors by dividing by $\sqrt{n}$ and denote the normalized vectors by $v_1, \ldots, v_{n+\delta n/2}$, $w_1, \ldots, w_{\delta n/2}$, that is, $v_j = \varepsilon_j/\sqrt{n}$ for $j = 1, \ldots, n+\delta n/2$ and $w_j = \varepsilon_{(n+\delta n/2+j)}/\sqrt{n}$ for $j = 1, \ldots, \delta n/2$. Throughout the proof c, c_1, c_2', C_3 etc. will denote universal constants which can be easily estimated.

Our proof mimics the proof of the theorem when the first n vectors form an orthonormal basis, and then the upper square in the matrix is an isometry. To substitute this fact, we will first of all need an estimate for the smallest eigenvalue of an $n \times (1+\delta/2)n$ matrix of random signs, which is given in Proposition 10 below, which is simply a reformulation of Theorem 2. It can be looked upon as a near-orthogonality result for the n random column vectors which are sign-vectors that live in $(n+\delta n/2)-$dimensional space.

Proposition 10. *There exist universal constants δ_0, c'', c_1 and c_1' such that for any $c''n^{-1/6}\log n < \delta < \delta_0$, if v_j are $n+\delta n/2$ random vectors chosen uniformly and independently in $\{-1/\sqrt{n}, 1/\sqrt{n}\}^n$ then with probability greater than $1-e^{-c_1'\delta n^{1/6}}$ we have for every $x \in \mathbb{R}^n$ that*

$$c_1\delta|x| \le \left(\sum_{j=1}^{n+\delta n} |\langle x, v_j\rangle|^2\right)^{1/2}.$$

The idea of the proof of Theorem 9 is to use the "near orthogonality" of the first $n+\delta n/2$ row vectors to ensure a lower bound in most directions. For the directions which remain, we obtain a lower bound by using the last $\delta n/2$ rows. To this end we will use a net argument, and hence we also need an upper bound for the contribution of the last $\delta n/2$ rows. This is given by the following

Lemma 11. *There exist universal constants c_3' and C_3 such that for any $\delta > 0$ if w_j are $\delta n/2$ random vectors of $\pm 1/\sqrt{n}$ then with probability greater than $1-e^{-c_3' n}$ we have for every $x \in \mathbb{R}^n$ that*

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\delta n/2} |\langle x, w_j\rangle| \le C_3\sqrt{\delta}|x|. \tag{15}$$

(Notice that although for a single point, in expectation, we have (15) with δ instead of $\sqrt{\delta}$, for the probability to suffice for the whole net we need to allow deviation of order $\sqrt{\delta}$ from the expectation.)

Proof Bernstein inequality implies that for any $t > 1$

$$\mathbb{P}[\frac{2}{\delta n}\sum_{j=1}^{\delta n/2} |\langle x, w_j\rangle| \ge t\frac{|x|}{\sqrt{n}}] \le e^{-ct^2\delta n}$$

for a universal c. We pick a 1/2-net on the sphere S^{n-1} with cardinality 5^n and pick $t = \sqrt{\frac{2\ln 5}{c\delta}}$. Then with probability greater than $1-5^{-n}$ we have that for every element x in the net

$$\frac{1}{\sqrt{n}}\sum_{j=1}^{\delta n/2}|\langle x, w_j\rangle| \leq t\delta/2.$$

Successive approximation of any point on the sphere by points in the net and homogeneity of the inequality (15) completes the proof, where $C_3 = \sqrt{2\ln 5/c}$. □

We will also need a covering result of Schütt [13], about the covering number of the unit ball of ℓ_1^m by euclidean balls: There exists a universal constant C_5 such that for every $k < m$

$$N\left(\sqrt{m}B(\ell_1^m), C_4\sqrt{\frac{m}{k}\log\frac{m}{k}}B(\ell_2^m)\right) \leq e^k \tag{16}$$

where for two convex bodies K and T the number $N(K,T)$ denotes the minimal number of translates of T needed to cover K.

Proof of Theorem 9 We define

$$\Sigma_\gamma = \{x \in S^{n-1} : \frac{1}{\sqrt{n}}\sum_{j=1}^{n+\delta n/2}|\langle x, v_j\rangle| \leq \gamma\}$$

(notice that we only use v_j and not w_j). If a point on the sphere is not in Σ_γ then a lower bound γ holds for this point for the left hand side of (14). We denote by A the $(n+\delta n/2)\times n$ matrix with rows v_j, and for convenience denote $m = n + \delta n/2$.

We now use (16) to cover $\mathrm{Im}A \cap \sqrt{n}B(\ell_1^m)$ by $e^{c_5'\delta n}$ balls of radius $r = C_4\sqrt{(\frac{1+\delta}{c_5'\delta})\log(\frac{1+\delta}{c_5'\delta})}$ (where c_5' is a universal constant to be determined later).

We have used the fact that taking a section only reduces the covering number by euclidean balls. Denote by $y_j \in \mathbb{R}^m \cap \mathrm{Im}A$ the centers of this covering, and let $x_j \in \mathbb{R}^n$ be their pre-images, so that $Ax_j = y_j$. Since there are only $e^{c_5'\delta n}$ of them, we can use Chernoff inequality in the following way: For a suitably chosen universal c_5 the probability that for a single i we have $|\langle x_j, w_i\rangle| \geq 3c_5|x_j|/\sqrt{n}$ is greater than 1/2 (this is not difficult to prove, see for example [1]). Therefore, by Chernoff's theorem, the probability that for at least 1/3 of the indices $i = 1, \dots, \delta n/2$ we have that $|\langle x_j, w_i\rangle| \geq 6c_5|x_j|/\sqrt{n}$ is greater than $1 - e^{-2c_5'\delta n}$ (this is our definition of c_5', which is universal). We get that with probability $1 - e^{-c_5'\delta n}$ for every j we have

$$c_5\delta|x_j| \leq \frac{1}{\sqrt{n}}\sum_{i=1}^{\delta n/2}|\langle x_j, w_i\rangle|.$$

Let $x \in S^{n-1}$, and consider

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n+\delta n/2}|\langle x, v_i\rangle| + \frac{1}{\sqrt{n}}\sum_{i=1}^{\delta n/2}|\langle x, w_i\rangle| \tag{17}$$

(which is the same as the left hand side of (14) up to a factor $(1+\delta)$). Recall that if $x \in S^{n-1}$ and $x \notin \Sigma_\gamma$, we have a lower bound at least γ for (17). Otherwise, we have $Ax \in \gamma\sqrt{n}B(\ell_1^m)$ (and of course also $Ax \in \mathrm{Im}A$). Therefore, there is some index j with $|Ax - \gamma Ax_j| < \gamma r$, where we use absolute value $|\cdot|$ to denote the euclidean norm. This implies, using Proposition 10 (which holds with probability at least $1 - e^{-c_1'\delta n^{1/6}}$), that $|x - \gamma x_j| < \gamma r/(c_1\delta)$. In particular, $|x_j| > \frac{1}{\gamma} - r/(c_1\delta)$. By (15) we know that this implies that

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{i=1}^{\delta n/2}|\langle x,w_i\rangle| &\geq \frac{1}{\sqrt{n}}\sum_{i=1}^{\delta n/2}|\langle \gamma x_j,w_i\rangle| - \frac{1}{\sqrt{n}}\sum_{i=1}^{\delta n/2}|\langle x-\gamma x_j,w_i\rangle| \\ &\geq c_5\delta\gamma|x_j| - C_5\sqrt{\delta}\gamma r/(c_1\delta) \\ &\geq c_5\delta - r\gamma(1+C_5\sqrt{\delta})/(c_1\delta).\end{aligned}$$

This tells us we may choose $\gamma = \delta^2 c_5 c_1/(2r(1+C_3\sqrt{\delta}))$, and have a lower bound $c_5\delta/2$ for this set. For the other set we have lower bound γ, that is (remembering what was r), $c_0\delta^{5/2}/\log(1/\delta)^{1/2}$ (for c_0 a universal constant suitably chosen). The proof is complete. □

References

[1] S. Artstein-Avidan, O. Friedland and V. Milman, *Some Geometric Applications of Chernoff-type Estimates*, to appear in GAFA Lecture Notes.

[2] Z. D. Bai and Y. Q. Yin, *Limit of the smallest eigenvalue of a large dimensional sample covariance matrix*, Ann. Probab. 21 (1993), no. 3, 1275–1294.

[3] A. Garnaev and E. Gluskin, *The widths of a Euclidean ball (in Russian)* Dokl. Akad. Nauk SSSR, 277(5), (1984), 1048-1052.

[4] W. Johnson and G. Schechtman, *Very tight embeddings of subspaces of L_p $1 \leq p < 2$ into ℓ_p^n*, Geom. Funct. Anal. 13 (2003), no. 4, 845–851.

[5] B. Kashin *Section of some finite-dimensional sets and classes of smooth functions (in Russian)* Izv. Acad. Nauk. SSSR 41 (1977) 334–351.

[6] T. Kusalik, J. A. Mingo, R. Speicher, *Orthogonal polynomials and fluctuations of random matrices*, preprint: arXiv:math.OA/0503169

[7] A. E. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, *Smallest singular value of random matrices and geometry of random polytopes*, Adv. in Math. 195, iss. 2, 491–523

[8] A. 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.

[9] V. A. Marchenko, L. A. Pastur, *Distribution of eigenvalues in certain sets of random matrices* (Russian), Mat. Sb. (N.S.) 72 (114), 1967, 507–536

[10] V. Milman and A. Pajor, *Regularization of star bodies by random hyperplane cut off*, Studia Mathematica 159 (2), 2003, 247–261.

[11] G. Schechtman, *Random embeddings of Euclidean spaces in sequence spaces*, Israel Journal of Mathematics 40, No. 2, 1981, 187–192.

[12] G. Schechtman, *Special orthogonal splittings of* L_1^{2k}, Israel Journal of Mathematics 139 (2004), 337–347.

[13] C. Schütt, *Entropy numbers of diagonal operators between symmetric Banach spaces*, J. Approx. Theory 40 (1984) no. 2, 121–1128

[14] S. Sodin, *On the spectrum of random covariance matrices*, in preparation

[15] S. Szarek, *On the best constants in Khinchin inequality*, Studia Math. 58 (1976) no. 2, 197–208.

Shiri Artstein-Avidan, Department of Mathematics, Princeton University, Fine Hall, Washington Road, Princeton NJ 08544-1000 USA and School of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton NJ 08540 USA.

Email address: artstein@princeton.edu

Omer Friedland, School of Mathematical Science, Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel.

Email address: omerfrie@post.tau.ac.il

Vitali Milman, School of Mathematical Science, Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel.

Email address: milman@post.tau.ac.il

Sasha Sodin, School of Mathematical Science, Tel Aviv University, Ramat Aviv, Tel Aviv, 69978, Israel.

Email address: sodinale@post.tau.ac.il