ON THE LARGEST AND THE SMALLEST SINGULAR VALUE OF SPARSE RECTANGULAR RANDOM MATRICES

F. GÖTZE* AND A. TIKHOMIROV

ABSTRACT. We derive estimates for the largest and smallest singular values of sparse rectangular $N \times n$ random matrices, assuming $\lim_{N,n\to\infty} \frac{n}{N} = y \in (0,1)$. We consider a model with sparsity parameter p_N such that $Np_N \sim \log^{\alpha} N$ for some $\alpha > 1$, and assume that the moments of the matrix elements satisfy the condition $\mathbb{E}\,|X_{jk}|^{4+\delta} \le C < \infty$. We assume also that the entries of matrices we consider are truncated at the level $(Np_N)^{\frac{1}{2}-\varkappa}$ with $\varkappa := \frac{\delta}{2(4+\delta)}$.

1. INTRODUCTION

In the last five to ten years, significant progress has been made in studying the asymptotic behavior of the spectrum of sparse random matrices. A typical example of such matrices is the incidence matrix of a random graph. Thus, for Bernoulli matrices Konstantin Tikhomirov obtained exact asymptotics for the probability of singularity, see [14]; also, see [9]. For the adjacency matrix of Erdös - Renyi random graphs, H.-T. Yau and L. Erdös & Co. proved a local semicircular law and investigated the behavior of the largest and the smallest singular values and as well as eigenvector statistics, see the papers of [2, 4] and the literature therein. In particular for adjacency matrices of regular graphs, local limit theorems and the behavior of extremal eigenvalues were investigated by H.-T. Yau and co-authors [1]. For non-Hermitian sparse random matrices M. Rudelson and K. Tikhomirov proved the circular law under unimprovable conditions on the probability of sparsity and the moments of distributions of the matrix elements (see [12]). J.O. Lee and J.Y. Hwang studied the spectral properties of sparse sample covariance matrices, which includes adjacency matrices of the bipartite Erdös–Renyi graph model). In [7] the authors prove a local law for the eigenvalues density up to the upper spectral edge assuming that sparsity probability p has order $N^{-1+\varepsilon}$ for some $\varepsilon > 0$ (here N denotes the growing order of the matrix) and entries of matrix X_{ij} are i.i.d. r.v.'s such that (in our notations)

$$\mathbb{E}\,|X_{11}|^2 = 1 \text{ and } \mathbb{E}\,|X_{11}|^q \le (Cq)^{cq} \text{ for every } q \ge 1. \tag{1.1}$$

They also prove the Tracy-Widom limit law for the largest eigenvalues of sparse sample covariance matrices. However, in the proof of the local Marchenko-Pastur law and the Tracy-Widom limit, they assume a priori that the result of [3, Lemma 3.11] holds for

---

Date: January 1, 2023.

Key words and phrases. Random matrices, sample covariance matrices, Marchenko–Pastur law.

Research funded by the German Research Foundation – SFB 1283/2 2021 – 317210226.

sparse matrices (see [7, Proposition 2.13]), which includes, in particular, the boundedness of the largest singular value that is the operator norm) of a sparse matrix. They don't investigate the smallest singular value of sparse rectangular matrices though.

We derive bounds for the *smallest* and the *largest* singular values of sparse rectangular random matrices assuming that the probability p_N decreases in such a way that $Np_N \geq \log^{\frac{2}{\varkappa}} N$ for some $\varkappa > 0$, and that the moment conditions are weaker than those in (1.1) (see condition (1.6)). Our main result is devoted to the smallest singular value of a sparse rectangular random matrix from an ensemble of *dilute* Wigner type matrices.

Suppose $n \geq 1$ and $N > n$. Consider independent identically distributed zero mean random variables X_{jk}, $1 \leq j \leq N$, $1 \leq k \leq n$ with $\mathbb{E}\, X_{jk}^2 = 1$ (where the distribution of X_{jk} may depend on N), which are independent of a set of independent Bernoulli random variables ξ_{jk}, $1 \leq j \leq N$, $1 \leq k \leq n$, with $\mathbb{E}\, \xi_{jk} = p_N$. In what follows we shall simplify notation by denoting $p = p_N$. We now introduce the following model of dilute sparse matrices as a sequence of random matrices of the following type

$$\mathbf{X} = (\xi_{jk} X_{jk})_{1 \leq j \leq N, 1 \leq k \leq n}. \tag{1.2}$$

Denote by $s_1 \geq \cdots \geq s_n$ the singular values of $\mathbf{X}$, and let $\mathbf{Y} = \mathbf{X}^* \mathbf{X}$ denote the sample covariance matrix.

Put $y = y(N, n) = \frac{n}{N}$. We shall assume that $y(N, n) \to y_0 < 1$ as $N, n \to \infty$. In what follows we shall vary the parameter N only.

Theorem 1.1. *Let $\mathbb{E}\, X_{jk} = 0$ and $\mathbb{E}\, |X_{jk}|^2 = 1$. Suppose that there exists a positive constant $C > 0$ such that*

$$\mathbb{E}\, |X_{jk}|^{4+\delta} \leq C < \infty, \tag{1.3}$$

for any $j, k \geq 1$ and for some $\delta > 0$. Suppose also that there exists a positive constant B, such that

$$Np \geq B \log^{\frac{3}{2\varkappa}} N, \tag{1.4}$$

where $\varkappa = \frac{\delta}{2(4+\delta)}$.

Then for every $Q \geq 1$ and $A > 0$ there exists a constant $K = C(Q, \delta, \mu_{4+\delta}, A, B)$ such that

$$\Pr\{\, s_1 \geq K\sqrt{Np}\} \leq CN^{-Q} + N^2 p \Pr\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa} \ln N\}.$$

Theorem 1.2. *Let $\mathbb{E}\, X_{jk} = 0$ and $\mathbb{E}\, |X_{jk}|^2 = 1$. Suppose that*

$$\mathbb{E}\, |X_{11}|^4 = \mu_4 < \infty,$$

and there exists a positive constant B, such that

$$Np \geq B \log^2 N. \tag{1.5}$$

Then there exists a constant $\tau_0 > 0$ such that for every $\tau \leq \tau_0$, $Q \geq 1$ and $K > 0$ there exists a constant $C = C(Q, \mu_4, K, B)$ with

$$\Pr\{\, s_n \leq \tau\sqrt{Np}\} \leq CN^{-Q} + \Pr\{s_1 > K\sqrt{Np}\}.$$

These results immediately imply the following corollary.

Corollary 1.3. *Under conditions of Theorem 1.1 there exist a constant $\tau_0 > 0$ such that for any $\tau \le \tau_0$ and for any $A > 0$ there exists a constant $C = C(A, \delta)$ depending on A and δ such that the following inequality holds*

$$\Pr\{\, s_n \le \tau\sqrt{Np}\} \le CN^{-Q} + N^2 p \Pr\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\}.$$

Corollary 1.4. *Assume the conditions of Theorem 1.1. In addition assume that there exists a constant B such that for every $N \ge 1$*

$$p = p_N \ge B/\ln^4 N.$$

Then

$$\Pr\{\, s_1 \ge K\sqrt{Np}\} \le CN^{-Q} + \frac{C}{\ln^{\delta} N}.$$

Proof. Applying Markov's inequality, we obtain

$$\Pr\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\} \le \frac{\mu_{4+\delta}}{(Np)^2\ln^{4+\delta} N}.$$

By the conditions of Corollary 1.4, we get

$$\Pr\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\} \le \frac{\mu_{4+\delta}}{N^2 B^{4+\delta}\ln^{\delta} N}.$$

The result follows now immediately from theorem 1.1. Thus, Corollary 1.4 is proved. □

We may consider random variables X_{ij} for $i = 1, \dots, N; j = 1, \dots, n$, with identical distributions depending on N. In this case we have the following result.

Corollary 1.5. *In addition to conditions of Theorem 1.1 assume that for any q such that $4 + \delta \le q \le C \log n$*

$$\mathbb{E}\,|X_{11}|^q \le C_0^q q^q (Np)^{q(\frac{1}{2}-\varkappa)-2}. \tag{1.6}$$

Then for every $Q \ge 1$ and $A > 0$ there exist constants $K = K(Q, \delta, \mu_{4+\delta}, A)$ and $C = C(Q, \delta, \mu_{4+\delta}, A)$ such that

$$\Pr\{\, s_1 \ge K\sqrt{Np}\} \le CN^{-Q}.$$

and there exists a constant $\tau_0 > 0$ such that for every $\tau \le \tau_0$, $Q \ge 1$ there exists a constant $C = C(Q, \delta, \mu_{4+\delta})$

$$\Pr\{\, s_n \le \tau\sqrt{Np}\} \le CN^{-Q} \tag{1.7}$$

2. Proof of Theorem 1.1

Let $\widetilde{X}_{ij}$ denote truncated random variables X_{ij}, i.e.

$$\widetilde{X}_{ij} = X_{ij}\mathbb{I}\{|X_{ij}| \le A(Np)^{\frac{1}{2}-\varkappa}\ln N\},$$

where $\mathbb{I}\{B\}$ denotes the indicator of an event B. Let $\widetilde{\mathbf{X}}$ denote the matrix with entries $\xi_{ij}\widetilde{X}_{ij}$. By $\|\mathbf{A}\|$ we denote the operator norm of a matrix $\mathbf{A}$. First we estimate the spectral norm of the matrix $\mathbb{E}\,\widetilde{\mathbf{X}}$. Since X_{ij} and ξ_{ij} are identically distributed random variables we have

$$\|\,\mathbb{E}\,\widetilde{\mathbf{X}}\| = np|\,\mathbb{E}\,\widetilde{X}_{11}|.$$

By condition (1.3), we have

$$|\mathbb{E}\,\widetilde{X}_{11}| = |\mathbb{E}\,X_{11}\mathbb{I}\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\}| \le \frac{C}{A^3(Np)^{\frac{3}{2}+\varkappa}}.$$

From here we get the bound

$$\|\,\mathbb{E}\,\widetilde{\mathbf{X}}\| \le CA^{-3}(Np)^{-\frac{1}{2}-\varkappa}. \tag{2.1}$$

We consider now the centered and truncated random variables $\widehat{X}_{ij} = \widetilde{X}_{ij} - \mathbb{E}\,\widetilde{X}_{ij}$ for $i = 1,\dots N, j = 1,\dots n$, and the matrix $\widehat{\mathbf{X}} = (\xi_{ij}\widehat{X}_{ij})$. Let $\widehat{s}_1 \ge \widehat{s}_2 \ldots \ge \widehat{s}_n$ denote the singular values of the matrix $\widehat{\mathbf{X}}$ and resp. let $\widetilde{s}_1 \ge \widetilde{s}_2 \ldots \ge \widetilde{s}_n$ denote the singular values of the matrix $\widetilde{\mathbf{X}}$. Note that

$$\begin{aligned}\Pr\{s_1 \ne \widetilde{s}_1\} \le \Pr\{\mathbf{X} \ne \widetilde{\mathbf{X}}\} &\le \sum_{i=1}^N\sum_{j=1}^n p\Pr\{\widetilde{X}_{ij} \ne X_{ij}\}\\ &= nNp\Pr\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\}\end{aligned} \tag{2.2}$$

Furthermore, we have

$$\widetilde{s}_1 \le \widehat{s}_1 + \|\,\mathbb{E}\,\widetilde{\mathbf{X}}\|. \tag{2.3}$$

According to (2.1) we may assume that

$$\|\,\mathbb{E}\,\widetilde{\mathbf{X}}\| \le \gamma\sqrt{Np} \tag{2.4}$$

for sufficiently small $\gamma > 0$. We may write now

$$\Pr\{s_1 > K\sqrt{Np}\} \le \Pr\{\widehat{s}_1 > \frac{1}{2}K\sqrt{Np}\} + N^2p\Pr\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\} \tag{2.5}$$

Note that

$$\begin{aligned}\widehat{\sigma}_n^2 &= \mathbb{E}\,\widehat{X}_{11}^2 = \mathbb{E}(\widetilde{X}_{11})^2 - (\mathbb{E}\,\widetilde{X}_{11})^2\\ &= 1 - \mathbb{E}\,X_{11}^2\mathbb{I}\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\} - (\mathbb{E}\,X_{11}\mathbb{I}\{|X_{11}| > A(Np)^{\frac{1}{2}-\varkappa}\ln N\})^2.\end{aligned} \tag{2.6}$$

It is easy that

$$|1-\sigma_n| \le |1-\sigma_n^2| \le \frac{2\mu_{4+\delta}}{A^{2+\delta}(Np)^{(2+\delta)(\frac{1}{2}-\varkappa)}}. \tag{2.7}$$

Without loss of generality we may assume that $\sigma_n \ge \frac{1}{2}$. Consider now the matrix $\breve{\mathbf{X}} = \frac{1}{\sigma_n}\widehat{\mathbf{X}}$. Let $\breve{s}_1$ denote the largest singular value of the matrix $\breve{\mathbf{X}}$. Then

$$\Pr\{\widehat{s}_1 > K\sqrt{Np}\} \le \Pr\{\breve{s}_1 > 2K\sqrt{Np}\}. \tag{2.8}$$

During the rest of the proof of Theorem 1.1 we shall consider the matrix $\mathbf{X}$ with entries $\xi_{ij}X_{ij}$, $i = 1,\dots,N$ $j = 1,\dots,n$ satisfying the following conditions (CI):

- ξ_{ij} are independent Bernoulli r.v.'s with $\mathbb{E}\,\xi_{ij} = p\,(= p_N)$;
- X_{ij} are i.i.d. r.v.'s for $1 \le i \le N, 1 \le j \le n$, such that $\mathbb{E}\,X_{11} = 0$, $\mathbb{E}\,|X_{11}|^{4+\delta} \le \mu_{4+\delta}$ and
$$|X_{11}| \le A(Np)^{\frac{1}{2}-\varkappa}\ln N \qquad \text{a.s.}$$

We use the following result of Seginer (see [13, Corollary 2.2]).

Proposition 2.1. *There exists a constant A such that for any $N, n \geq 1$, any $q \leq 2\log\max\{n, N\}$, and any $N \times n$ random matrix $\mathbf{X} = (X_{ij})$ where X_{ij} are i.i.d. zero mean random variables, the following inequality holds:*

$$\max\Big\{\mathbb{E}\max_{1\leq i\leq N}\|\mathbf{X}_{i\cdot}\|_2^q, \mathbb{E}\max_{1\leq j\leq n}\|\mathbf{X}_{\cdot j}\|_2^q\Big\} \leq \mathbb{E}\|\mathbf{X}\|^q$$
$$\leq (2A)^q\Big(\mathbb{E}\max_{1\leq i\leq N}\mathbb{E}\|\mathbf{X}_{i\cdot}\|_2^q + \max_{1\leq j\leq n}\|\mathbf{X}_{\cdot j}\|_2^q\Big). \tag{2.9}$$

Here $\mathbf{X}_{i\cdot}$, resp. $\mathbf{X}_{\cdot j}$, denote the i-th row, resp. the j-th column of $\mathbf{X}$.

Proof of Theorem 1.1. Note that $s_1 = \|\mathbf{X}\|$. Using the notations introduced above, we now estimate $\mathbb{E}\|\mathbf{X}_{i\cdot}\|^q$. By the definition of $\mathbf{X}$ we have

$$\mathbb{E}\|\mathbf{X}_{i\cdot}\|_2^q = \mathbb{E}\Big(\sum_{k=1}^n X_{ik}^2\xi_{ik}\Big)^{\frac{q}{2}} \leq 2^{q-1}\Big(\sum_{k=1}^n \mathbb{E}\,X_{ik}^2\xi_{ik}\Big)^{\frac{q}{2}} + 2^{q-1}\,\mathbb{E}\,\Big|\sum_{k=1}^n (X_{ik}^2-1)\xi_{ik}\Big|^{\frac{q}{2}}. \tag{2.10}$$

Note that

$$\mathbb{E}\,X_{ik}^2\xi_{ik} = p. \tag{2.11}$$

Now, applying Rosenthal's inequality we get

$$\mathbb{E}\Big|\sum_{k=1}^n (X_{ik}^2-1)\xi_{ik}\Big|^{\frac{q}{2}} \leq C^q\Big(q^{\frac{q}{4}}\Big(\sum_{k=1}^n \mathbb{E}(X_{ik}^2-1)^2\xi_{ik}\Big)^{\frac{q}{4}} + q^{\frac{q}{2}}p\sum_{k=1}^n \mathbb{E}\,|X_{ik}^2-1|^{\frac{q}{2}}\Big), \tag{2.12}$$

which implies

$$\mathbb{E}\Big|\sum_{k=1}^n (X_{ik}^2-1)\xi_{ik}\Big|^{\frac{q}{2}} \leq C^q\big(q^{\frac{q}{4}}(Np)^{\frac{q}{4}} + q^{\frac{q}{2}}Np\,\mathbb{E}\,|X_{11}|^q\big). \tag{2.13}$$

By assumptions (CI), we have

$$\mathbb{E}\,|X_{11}|^q \leq C^q(Np)^{\frac{q}{2}-q\varkappa-2}\ln^{q-4-\delta} N. \tag{2.14}$$

Note that for $q \sim \ln N$ inequality (2.14) coincide with condition (1.6). Combining inequalities (2.10)–(2.14), we now get

$$\mathbb{E}\|\mathbf{X}_{i\cdot}\|_2^q \leq C^q(Np)^{\frac{q}{2}}\Big(1 + \Big(\frac{q}{Np}\Big)^{\frac{q}{4}} + N^{-1}p^{-1}\ln^{-(4+\delta)} N\,\Big(\frac{q\ln^2 N}{(Np)^{2\varkappa}}\Big)^{\frac{q}{2}}\Big).$$

Taking into account (1.5), as well as $q \leq C\log n$, we obtain, for $q \leq 2\log\max\{n, N\}$,

$$\mathbb{E}\|\mathbf{X}_{i\cdot}\|_2^q \leq C^q(Np)^{\frac{q}{2}}.$$

A similar bound holds for $\mathbb{E}\|\mathbf{X}_{\cdot j}\|^q$. We may now write

$$\mathbb{E}\|\mathbf{X}\|^q \leq C^qN(Np)^{\frac{q}{2}}.$$

Taking $K \gg C$ and applying Markov's inequality, the claim follows. Thus Theorem 1.1 is proved. $\square$

3. Smallest singular values

We shall now prove Theorem 1.2 using an approach developed by Litvak, Pajor, Rudelson [8], Rudelson and Vershynin in [10] for rectangular matrices for the case $p = 1$ and Götze and Tikhomirov in [5] for the sparse dilute Wigner matrices. Denote by $\mathcal{S}^{(n-1)}$ the unit sphere in $\mathbb{R}^n$. Let $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{S}^{(n-1)}$ be a fixed unit vector and $\mathbf{X}$ be a matrix defined in (1.2).

We divide the vectors on the sphere into two parts: compressible and incompressible vectors recalling the definition.

Definition 3.1. Let $\delta, \rho \in (0,1)$. A vector $\mathbf{x} \in \mathbb{R}^n$ is called sparse if $|\mathrm{supp}(\mathbf{x})| \le \delta n$. A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called compressible if $\mathbf{x}$ is within Euclidean distance ρ from the set of all sparse vectors. A vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ is called incompressible if it is not compressible. The sets of compressible and incompressible vectors will be denoted by $\mathit{Comp}(\delta,\rho)$ and $\mathit{Incomp}(\delta,\rho)$.

Note that

$$s_n = \inf_{\mathbf{x} \in \mathcal{S}^{(n-1)}} \|\mathbf{X}\mathbf{x}\|_2$$

and

$$\Pr\{s_n \le \tau\sqrt{Np}\} \le \Pr\{\inf_{x \in \mathit{Comp}(\delta,\rho)} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\} + \Pr\{\inf_{x \in \mathit{Incomp}(\delta,\rho)} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\}, \tag{3.1}$$

for some $\delta, \rho \in (0,1)$ and $\tau > 0$, not depending on n.

For sparse matrices with $p = p_N \to 0$ as $N \to \infty$ we cannot directly estimate the first term on the right hand side of (3.1) using the well-known two step approach of estimating $\Pr\{\|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\}$ for a fixed vector $\mathbf{x} \in \mathcal{S}^{(n-1)}$ followed by a union bound for the some ε-net of $\mathit{Comp}(\delta,\rho)$ and arriving at a bound for the infimum of $\mathbf{x} \in \mathit{Comp}(\delta_n,\rho)$ with $\delta_n \sim p$ going to zero. The Rudelson - Vershynin methods for incompressible vectors won't work in this case. In order to estimate $\Pr\{\inf_{x \in \mathit{Comp}(\delta,\rho)} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\}$ with some $\delta > 0$ which does not not depend on n, we shall use a method developed in Götze-Tikhomirov [5]. This is based on a recurrence approach which allows us to increase δ_N step by step Np times arriving in $\log N$ steps at an estimate of $\delta > \delta_0$ which does not depend on N. The details of this approach will be described in Section 3.1.

In Section 3.3 we shall derive bounds for $\Pr\{\inf_{x \in \mathit{Incomp}(\delta,\rho)} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\}$.

3.1. Compressible vectors.

Let L be an integer such that

$$\left(\frac{\delta_0 Np}{|\log p| + 1}\right)^{L-1} \le p^{-1} \le \left(\frac{\delta_0 Np}{|\log p| + 1}\right)^{L}, \tag{3.2}$$

where $\delta_0 \in (0,1)$ denotes some constant independent on N. Note that under the conditions of Theorem 1.2

$$L \le c \log N / \log\log N \tag{3.3}$$

with a constant $c = c(\delta_0)$. We introduce a set of numbers $p_{\nu N}$ and $\delta_{\nu N}$, for $\nu = 1, \ldots, L$, as follows

$$p_{\nu N} = (Np)\delta_{\nu-1N} \text{ and } \delta_{\nu N} = \delta_0 p_{\nu N}/(1 + |\log p_{\nu N}|).$$

Here

$$p_{0N} = p \text{ and } \delta_{0N} = \delta_0 p/(1 + |\log p|).$$

Furthermore, introduce as well

$$\widehat{p}_{\nu N} = \left(\frac{Np\delta_0}{|\log p| + 1}\right)^{\nu} p \text{ and } \widehat{\delta}_{\nu N} := \left(\frac{\delta_0 Np}{|\log p| + 1}\right)^{\nu-1} \frac{\delta_0 p}{|\log p| + 1}.$$

Lemma 3.2. *The following inequalities hold*

$$p_{\nu,N} \geq \widehat{p}_{\nu} \tag{3.4}$$

and

$$\delta_{\nu,N} \geq \widehat{\delta}_{\nu,N}, \tag{3.5}$$

for $\nu = 1, \dots, N$

Proof. By condition of Theorem 1.2,

$$\frac{Np}{1 + |\ln p|} \geq B \ln N. \tag{3.6}$$

Without loss of generality we may assume that

$$\frac{Np\delta_0}{1 + |\ln p|} > 1. \tag{3.7}$$

It is straightforward to check now that $p_{\nu,N} \geq p$, for $\nu = 1, \dots, N$. In fact, for $\nu = 1$ it is easy. Assume that for some $\nu = 1, \dots, N-1$ the inequality $p_{\nu-1,N} \geq p$ holds. Then

$$p_{\nu,N} = \frac{Np\delta_0 p_{\nu-1,N}}{1 + |\ln p_{n-1,N}|} \geq \frac{Np\delta_0}{1 + |\ln p|} p_{\nu-1,N} \geq \frac{Np\delta_0}{1 + |\ln p|} p \geq p. \tag{3.8}$$

We may write now the following inequalities

$$\delta_{\nu,N} \geq \frac{\delta_0}{1 + |\ln p|} p_{\nu,N} \tag{3.9}$$

and

$$p_{\nu,N} \geq \frac{Np\delta_0}{1 + |\ln p|} p_{\nu-1,N}, \tag{3.10}$$

for $\nu = 1, \dots, N$. Applying induction for the last inequality, we get, for $\nu = 1, \dots, N$,

$$p_{\nu,N} \geq \widehat{p}_{\nu,N}. \tag{3.11}$$

The last inequality implies that, for $\nu = 1, \dots, N$,

$$\delta_{\nu,N} \geq \frac{\delta_0}{1 + |\ln p|} \widehat{p}_{\nu-1,N} = \left(\frac{Np\delta_0}{1 + |\ln p|}\right)^{\nu-1} \frac{p\delta_0}{1 + |\ln p|} = \widehat{\delta}_{nu,N}. \tag{3.12}$$

Thus, lemma is proved. □

Corollary 3.3. *There exist constants* $\gamma_0 > 0, \gamma_1 > 0$ *such that*

$$\delta_{L,N} \geq \gamma_0 \text{ and } p_{LN} \geq \gamma_1. \tag{3.13}$$

Introduce the sets

$$\mathcal{C}_\nu := \mathrm{Comp}(\delta_{\nu,N}, \rho), \quad \mathcal{IC}_\nu := \mathrm{Incomp}(\delta_{\nu,N}, \rho), \quad \nu = 0, \ldots, L.$$

Note that $L \geq 1$ for $Np^2/(|\log p| + 1) \leq D$ with some constant D. The case $Np^2/(|\log p| + 1) \geq D$ will we treated separately. In what follows we shall assume that $L \geq 1$.

Definition 3.4. The Lévy concentration function of a random variable ξ is defined for $\varepsilon > 0$ as

$$\mathcal{L}(\xi, \varepsilon) = \sup_{v \in \mathbb{R}} \Pr\{|\xi - v| \leq \varepsilon\}. \tag{3.14}$$

By $\mathbf{P}_{\mathbb{E}}$ we denote the orthogonal projection in $\mathbb{R}^n$ onto a subspace $\mathbb{E}$. Similarly, by $\mathbf{P}_{\mathbb{J}}$ we denote the orthogonal projection onto $\mathbb{R}^{\mathbb{J}}$, where $\mathbb{J} \subset \{1, 2, \ldots, n\}$.

We reformulate and prove some auxiliary results from [10] below for our sparsity model.

First we prove an analog of [10, Lemma 3.2].

Lemma 3.5. *Let* $\mathbf{x} \in \mathcal{IC}_\nu$, $\nu = 1, \ldots, L$. *Let*

$$\zeta_j = \sum_{k=1}^n x_k \xi_{jk} X_{jk}, \quad j = 1, \ldots, N.$$

Then there exists some absolute constant A *such that*

$$\mathcal{L}(\frac{1}{\sqrt{p}}\zeta_j, \frac{\rho}{2}) \leq 1 - A\rho^4 p_{\nu N}. \tag{3.15}$$

Remark 3.6. *For* $\nu = L$ *there exists some constant* $0 < b < 1$ *such that*

$$\mathcal{L}(\frac{1}{\sqrt{p}}\zeta_j, \frac{\rho}{2}) \leq 1 - b < 1.$$

Proof. By Lemma 3.11 there exists a set $\sigma(x)$ such that for $k \in \sigma(\mathbf{x})$

$$\frac{1}{2\sqrt{n}} \leq |x_k| \leq \frac{1}{\sqrt{2n\delta_{\nu-1,N}}}, \text{ and } \|\mathbf{P}_{\sigma(\mathbf{x})}\mathbf{x}\|_2^2 \geq \rho^2.$$

Let

$$\eta = \sum_{k \in \sigma(\mathbf{x})} x_k \xi_{jk} X_{jk}/\sqrt{p}.$$

Note that

$$\mathbb{E}\,\eta^2 \geq \rho^2, \quad \mathbb{E}\,|\eta|^4 \leq A_0(1 + \frac{1}{N\delta_{\nu-1,N}p}).$$

Without loss of generality we may assume that $N\delta_{\nu-1,N}p \leq 1$. This implies that

$$\mathbb{E}\,|\eta|^4 \leq \frac{2A_0}{N\delta_{\nu-1,N}p}. \tag{3.16}$$

Let $Z = \eta - v$. Note that

$$\mathbb{E}\,Z^2 = \mathbb{E}\,\eta^2 + v^2 \geq v^2 + \rho^2,$$

and

$$\mathbb{E}\,\eta^4 \geq (\mathbb{E}\,\eta^2)^2 \geq \rho^4.$$

Using Minkowski's inequality, we get

$$\mathbb{E}^{\frac14}|Z|^4 \le \mathbb{E}^{\frac14}|\eta|^4 + v \le \mathbb{E}^{\frac14}|\eta|^4(1+\frac{v}{\rho}) \le \rho^{-1}\sqrt{2}\,\mathbb{E}^{\frac14}|\eta|^4(\rho^2+v^2)^{\frac12}.$$

Using the Paley-Zygmund inequality, we get

$$\Pr\{|\eta - v| > \varepsilon\} \ge \frac{\rho^4(\mathbb{E}|Z|^2-\varepsilon^2)^2}{4\,\mathbb{E}|\eta|^4(\rho^2+v^2)^2} \ge \frac{1}{4\,\mathbb{E}|\eta|^4}\frac{\rho^4(\rho^2+v^2-\varepsilon^2)^2}{(\rho^2+v^2)^2}.$$

The last inequality and inequality (3.16) together imply

$$\Pr\{|\eta - v| \ge \varepsilon\} \ge A_1\rho^4 N\delta_{\nu-1,N}p(1-\frac{2\varepsilon^2}{\rho^2+v^2}).$$

Finally, we may write

$$\Pr\{|\eta - v| \ge \frac12\rho\} \ge \frac12 A_1\rho^4 p_{\nu,N}.$$

Thus Lemma 3.5 is proved. □

For the set of sparse vectors the following lemma holds.

Lemma 3.7. *The following inequality holds.*

$$\mathcal{L}(\xi X/\sqrt{p}, \frac12) \le 1 - \frac{p}{8\mu_4}$$

Proof. For the proof it is enough to note that by the Paley-Zygmund inequality we have

$$\Pr\{|\xi X - v| \ge \frac12\} \ge p\frac{1+v^2-\varepsilon^2}{4\,\mathbb{E}|X|^4(1+v^2)^2} \ge \frac{p}{8\mu_4}$$

□

Lemma 3.8. *Let $\zeta_1,\dots,\zeta_N$ denote independent identically distributed random variables such that*

$$\Pr\{|\zeta_j| \le \lambda_n\} \le 1 - q_N,$$

for some $\lambda_N > 0$ and $q_N \in (0,1)$. Then there exist constants c, C such that

$$\Pr\{\sum_{j=1}^N \zeta_j^2 \le CNq_N\lambda_N^2\} \le \exp\{-cNq_N\}. \tag{3.17}$$

For the proof of this lemma see [5, Lemma 4.5].

We start with the estimation of $\|\mathbf{X}\mathbf{x}\|_2$ for a fixed $\mathbf{x} \in \mathcal{S}^{(n-1)}$.

Lemma 3.9. *There exist positive absolute constants τ_0 and c_0 such that*

$$\Pr\{\|\mathbf{X}\mathbf{x}\|_2 \le \tau_0\sqrt{Np}\} \le \exp\{-c_0Np\}.$$

Proof of Lemma 3.9. The proof of this lemma may be found in [5, Lemma 4.1], but for readers convenience we repeat it here. Let

$$\zeta_j = \sum_{k=1}^n X_{jk}\xi_{jk}x_k, \quad j = 1,\dots,N$$

Then

$$\|\mathbf{Xx}\|_2^2 = \sum_{j=1}^N \zeta_j^2.$$

Furthermore, we may write for $\tau > 0$ and any t

$$\Pr\{\sum_{j=1}^N \zeta_j^2 \le \tau^2 Np\} = \Pr\{\frac{\tau^2 Np}{2} - \frac{1}{2}\sum_{j=1}^N \zeta_j^2 \ge 0\} \le \exp\{Np\tau^2 t^2/2\} \prod_{j=1}^N \mathbb{E}\exp\{-t^2\zeta_j^2/2\}.$$

Using $\mathrm{e}^{-t^2/2} = \mathbb{E}\,\mathrm{e}^{it\eta}$, where η is a standard Gaussian random variable, we obtain

$$\Pr\{\sum_{j=1}^N \zeta_j^2 < \tau^2 np\} \le \exp\{Np\tau^2 t^2/2\} \prod_{j=1}^N \mathbb{E}_{\eta_j} \prod_{k=1}^n \mathbb{E}_{\xi_{jk}X_{jk}} \exp\{it\xi_{jk}X_{jk}x_k\eta_j\}, \quad (3.18)$$

where η_j, $j = 1, \ldots, N$ denote i.i.d. Gaussian standard r.v.s and $\mathbb{E}_Z$ denotes expectation with respect to Z conditional on all other r.v.s.

Take $\alpha = \Pr\{|\eta_1| \le C_1\}$ for some absolute positive constant C_1 which will be chosen later. Then it follows from 3.18 that

$$\Pr\{\sum_{j=1}^N \zeta_j^2 < \tau^2 Np\} \le \exp\{t^2\tau^2 Np/2\}$$
$$\times \prod_{j=1}^N \Big(\alpha\Big|\,\mathbb{E}_{\eta_j}\Big\{\prod_{k=1}^n \mathbb{E}_{\xi_{jk}X_{jk}} \exp\{it\eta_j x_k X_{jk}\xi_{jk}\}\Big||\eta_j| \le C_1\Big\}\Big| + 1 - \alpha\Big).$$

Note that for any $\alpha, x \in [0,1]$, and $\beta \le \alpha$

$$1 - \alpha + \alpha x \le \max\{x^\beta, \Big(\frac{\beta}{\alpha}\Big)^{\frac{\beta}{1-\beta}}\}.$$

Furthermore, we have

$$|\,\mathbb{E}_{\xi_{jk}X_{jk}} \exp\{it\xi_{jk}X_{jk}x_k\eta_j\}| \le \exp\{-\frac{p}{2}(1 - |f_{jk}(tx_k\eta_j)|^2)\}, \quad (3.19)$$

where $f_{jk}(u) = \mathbb{E}\exp\{iuX_{jk}\}$. Choose a constant $M > 0$ such that

$$\sup_{j,k\ge 1} \mathbb{E}\,|X_{jk}|^2 \mathbb{I}\{|X_{jk}| > M\} \le \frac{1}{2}.$$

Since $1 - \cos x \ge \frac{11}{24}x^2$ for $|x| \le 1$, conditioning on the event $|\eta_j| \le C_1$, we get for $|t| \le \frac{1}{MC_1}$,

$$1 - |f_{jk}(tx_k\eta_j)|^2 = \mathbb{E}_{X_{kj}}(1 - \cos(tx_k\widetilde{X}_{kj}\eta_j) \ge \frac{11}{24}x_k^2 t^2\eta_j^2\,\mathbb{E}\,|\widetilde{X}_{kj}|^2\mathbb{I}\{|X_{kj}| \le M\}. \quad (3.20)$$

Here we denote by $\widetilde{X}_{kj}$ the symmetrization of the r.v. X_{kj}. It follows from (3.19) for $|t| \le 1/(MC_1)$, that for $|\eta_j| \le C_1$,

$$|\,\mathbb{E}_{\xi_{jk}X_{jk}} \exp\{it\xi_{jk}X_{jk}x_k\eta_j\}| \le \exp\{-cpt^2x_k^2\eta_j^2\} \quad (3.21)$$

This implies that

$$|\prod_{k=1}^{n}\mathbb{E}_{\xi_{kj}X_{kj}}\exp\{it\eta_j x_k\xi_{jk}X_{jk}\}|\le\exp\{-cpt^2\eta_j^2\}. \tag{3.22}$$

We may choose C_1 large enough such that following inequalities hold for $|t|\le 1/MC_1$:

$$|\,\mathbb{E}_{\eta_j}\{\exp\{-cpt^2\eta_j^2\}\big||\eta_j|\le C_1\}|\le\exp\{-ct^2p/24\}. \tag{3.23}$$

Then we obtain

$$\Pr\{\sum_{j=1}^{N}\zeta_j^2\le\tau^2Np\}\le\exp\{Np\tau^2t^2/2\}\Big(\exp\{-c\beta t^2Np/24\}+\Big(\frac{\beta}{\alpha}\Big)^{N\frac{\beta}{1-\beta}}\Big) \tag{3.24}$$

Furthermore, we may take C_1 sufficiently large such that $\alpha\ge\frac{4}{5}$ and choose $\beta=\frac{2}{5}$. We get

$$\Pr\{\sum_{j=1}^{N}\zeta_j^2\le\tau^2Np\}\le\exp\{Np\tau^2t^2/2\}\Big(\exp\{-ct^2Np/60\}+2^{-2N/3}\Big). \tag{3.25}$$

For $\tau<\min\{\frac{\sqrt{c}}{\sqrt{60}},\frac{\sqrt{\ln 2}}{\sqrt{3}}MC_1\}$, we have for $|t|\le 1/(MC_1)$,

$$\Pr\{\sum_{j=1}^{N}\zeta_j^2\le\tau^2Np\}\le\exp\{-ct^2Np/120\}. \tag{3.26}$$

This implies the claim. Thus the lemma is proved. □

3.2. **Compressible and Incompressible Vectors.** First we prove an analog of Lemma 2.6 from [10].

Lemma 3.10. *There exist positive absolute constants δ_0,τ_0,c_1 such that*

$$\Pr\{\inf_{\mathbf{x}\in Comp(\delta_{0N},\rho_0)}\|\mathbf{X}\mathbf{x}\|_2\le\tau_0\sqrt{Np},\quad\|\mathbf{X}\|\le K\sqrt{Np}\}\le\exp\{-c_1Np\},$$

where

$$\delta_{0N}=\delta_0p/(|\log p|+1),\quad\rho_0=\tau_0/2K. \tag{3.27}$$

Proof. Let $k=[n\delta_{0N}]$. Denote by $\mathcal{N}_\eta$ an η-net on the $\mathcal{S}^{(k-1)}\cap\mathbb{R}^k$. Choose $\eta=\tau_0/2K$ First we consider the set of all sparse vectors $Sparse(k)$ with support$(\mathbf{x})\le k$. Using Lemma 3.9 and a union bound, we get

$$\Pr\{\inf_{\mathbf{x}\in Sparse(\delta_{0N})}\|\mathbf{X}\mathbf{x}\|_2\le 2\rho_0\sqrt{np}\}\le\binom{n}{k}|\mathcal{N}_\eta|\exp\{-c_0Np\}.$$

Using Stirling's formula and Proposition 2.1 from [10], we get

$$\Pr\{\inf_{\mathbf{x}\in Sparse(\delta_{0N})}\|\mathbf{X}\mathbf{x}\|_2\le 2\tau_0\sqrt{Np}\}$$
$$\le\frac{4n\delta_{0N}}{\sqrt{2\pi n\delta_{0N}(1-\delta_{0N})}}\frac{(1+\frac{K}{\rho_0})^{n\delta_{0N}-1}}{\delta_{0N}^{n\delta_{0N}}(1-\delta_{0N})^{n(1-\delta_{0N})}}\exp\{-c_0Np\}.$$

Simple calculations show

$$\Pr\{\inf_{\mathbf{x}\in Sparse(\delta_{0N})}\|\mathbf{X}\mathbf{x}\|_2\le 2\tau_0\sqrt{Np}\}\le\sqrt{\frac{2n\delta_{0N}}{(1-\delta_{0N})\pi}}$$
$$\times\exp\{n\delta_{0N}\Big((1-\frac{1}{n\delta_{0N}})\frac{K}{\rho_0}-\log\delta_{0N}-(1-\delta_{0N})\frac{1}{\delta_{0N}}\log(1-\delta_{0N})\Big)-c_0Np\}.$$

If we choose

$$\delta_{0N}:=\delta_0p/(1+|\log p|)$$

for a sufficiently small absolute constant δ_0, we get

$$\Pr\{\inf_{\mathbf{x}\in Sparse(\delta_{0N})}\|\mathbf{X}\mathbf{x}\|_2\le 2\tau_0\sqrt{Np}\}\le\exp\{-c_1Np\}.$$

Thus the Lemma is proved. □

In what follows, we shall use a technique developed in Götze and Tikhomirov [5] which is based on the following lemmas.

Lemma 3.11. *Let $\rho,\delta\in(0,1)$. Assume that $\mathbf{x}\in Incomp(\delta,\rho)$. Then there exists a set $\sigma_0(x)$ such that $|\sigma_0(x)|\ge Cn\delta\rho^2$ and $\frac{1}{2\sqrt{n}}\le|x_k|\le\frac{1}{\sqrt{n\delta/2}}$ for $k\in\sigma_0(x)$, and*

$$\sum_{k\in\sigma_0(x)}|x_k|^2\ge\rho^2.$$

For a proof of this Lemma see for instance [11, Lemma 3.4].

Lemma 3.12. *Let $\mathbf{x}\in\mathcal{IC}_\nu$ for some $\nu=0,\dots,L-1$. Then there exist constants c_1 and c_2 such that for any $0<\tau\le\tau_0$*

$$\Pr\{\|\mathbf{X}\mathbf{x}\|_2\le\tau\sqrt{Np}\}\le\exp\{-c_1Np_{\nu+1N}\}.$$

Proof. We repeat the proof of Lemma 3.9 till (3.20).

Furthermore, by Lemma 3.11 there exists a set $\sigma_0(x)$ such that $\frac{1}{2\sqrt{n}}\le|x_k|\le\frac{1}{\sqrt{n\delta_{\nu N}/2}}$ for $k\in\sigma_0(x)$, and

$$\sum_{k\in\sigma_0(x)}|x_k|^2\ge\rho^2. \tag{3.28}$$

We may write now

$$\sum_{k=1}^n(1-|f(tx_kX_{jk}\eta_j)|^2)\ge\sum_{k\in\sigma_0(x)}(1-|f(tx_kX_{jk}\eta_j)|^2).$$

Note that for $k\in\sigma_0$, and for $|X_{jk}|\le M$, and for $|\eta_j|\le C$, we have

$$|tx_kX_{jk}\eta_j|\le\frac{|t|CM\sqrt{2}}{\sqrt{N\delta_{\nu N}}}.$$

Taking $t=\kappa\sqrt{N\delta_{\nu N}}$ for $\kappa=\frac{1}{CM\sqrt{2}}$, we get

$$|tx_kX_{jk}\eta_j|\le 1,$$

and

$$1-|f_{\eta_j}(tx_kX_{jk}\eta_j)|^2\ge\frac{11}{24}t^2x_k^2\eta_j^2\,\mathbb{E}\,|X_{jk}|^2\mathbb{I}\{|X_{jk}|\le M\}\ge\frac{11}{48}t^2x_k^2\eta_j^2.$$

Repeating now the last part of the proof of Lemma 3.9 and taking into account inequality (3.28), we obtain for $\tau<\rho\min\{\frac{\sqrt{c}}{\sqrt{60}},\frac{\sqrt{\ln 2}}{\sqrt{3}}MC_1\}$, and for $|t|=\kappa\sqrt{N\delta_{\nu N}}$,

$$|\prod_{k=1}^n\mathbb{E}_{\xi_{jk}X_{jk}}\exp\{it\eta_jx_k\xi_{jk}X_{jk}\}|\le\exp\{-c\rho^2pt^2\eta_j^2\},\tag{3.29}$$

where c is an absolute constant as in (3.22). We may choose C_1 large enough such that the following inequalities hold for $|t|=\kappa\sqrt{N\delta_{\nu N}}$:

$$|\,\mathbb{E}_{\eta_j}\{\exp\{-cpt^2\eta_j^2\}\big||\eta_j|\le C_1\}|\le\exp\{-ct^2p/24\}.\tag{3.30}$$

We use here that $|t|p\le\delta_0$ by (3.2). Then we obtain

$$\Pr\{\sum_{j=1}^n\zeta_j^2\le\tau^2Np\}\le\exp\{Np\tau^2t^2/2\}\Big(\exp\{-c\beta t^2Np/24\}+\Big(\frac{\beta}{\alpha}\Big)^{N\frac{\beta}{1-\beta}})\}\tag{3.31}$$

Furthermore, we may take C_1 large enough such that $\alpha\ge\frac45$ and choose $\beta=\frac25$. We get

$$\Pr\{\sum_{j=1}^n\zeta_j^2\le\tau^2Np\}\le\exp\{Np\tau^2t^2/2\}\Big(\exp\{-ct^2Np/60\}+2^{-2N/3}\Big).\tag{3.32}$$

For $\tau<\min\{\frac{\sqrt{c}}{\sqrt{60}},\frac{\sqrt{\ln 2}}{\sqrt{3}}MC_1\}$, we have for $|t|=\kappa\sqrt{N\delta_{\nu N}}$,

$$\Pr\{\sum_{j=1}^n\zeta_j^2\le\tau^2Np\}\le\exp\{-ct^2Np/120\}.\tag{3.33}$$

This inequality implies that

$$\Pr\{\sum_{j=1}^N\zeta_j^2\le\tau^2Np\}\le\exp\{-c(\rho^2N^2\kappa^2p\delta_{\nu N}\wedge N)/120\}.\tag{3.34}$$

Thus the lemma is proved. □

Furthermore, we consider the sets defined as

$$\widehat{\mathcal{C}}_\nu:=\mathcal{I}C_{\nu-1}\cap\mathcal{C}_\nu,\ \nu=1,\dots,L.\tag{3.35}$$

Lemma 3.13. *Under conditions of Theorem 1.2 we have, for $\nu=1,\dots,L$,*

$$\Pr\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_\nu}\|\mathbf{X}\mathbf{x}\|_2\le\tau\sqrt{Np}\}\le\exp\{-cNp_{\nu N}\}.$$

Proof. According to Lemma 3.12 we have for any fixed $\mathbf{x}\in\widehat{\mathcal{C}}_\nu$

$$\Pr\{\|\mathbf{X}\mathbf{x}\|_2\le2\tau\sqrt{Np}\}\le\exp\{-c_1Np_{\nu,N}\}.$$

Consider $\eta=\frac{\tau}{K}$-net $\mathcal{N}$ of $\widehat{\mathcal{C}}_\nu$. Then the event $\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_\nu}\|\mathbf{X}\mathbf{x}\|_2\le\tau\sqrt{Np}\}$ implies

$$\{\inf_{\mathbf{x}\in\mathcal{N}}\|\mathbf{X}\mathbf{x}\|_2\le2\tau\sqrt{Np}\}.\tag{3.36}$$

Without loss of generality we may assume that $\delta_{LN} < 1$. Using a union bound, we get

$$\Pr\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_\nu} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\} \le \binom{n}{n\delta_{\nu N}}|\mathcal{N}|\exp\{-c_1 N p_{\nu,N}\} \tag{3.37}$$

Using Stirling's formula and a simple bound for the cardinality of an η-net, for some sufficiently small absolute constant $\alpha_0 > 0$ (does not depend on ν) and

$$\delta_{\nu N} = \alpha_0 p_{\nu N}/(|\log p_{\nu,N}| + 1),\ p_{\nu N} := Np\delta_{\nu-1,N}$$

we get

$$\Pr\{\inf_{\mathbf{x}\in\widehat{\mathcal{C}}_\nu} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\} \le \exp\{-\widehat{c}_1 N p_{\nu N}\}.$$

Thus Lemma 3.13 is proved. □

Now we consider the case $Np^2/(|\log p| + 1) > D$ for some sufficiently large constant D. Let $\mathbf{x} \in \mathit{Incomp}(\delta_{0N}, \rho)$ and $\sigma(\mathbf{x})$ denote the set described in Lemma 3.11. Let

$$\zeta_j = \sum_{k=1}^n x_k \xi_{jk} X_{jk}, j = 1, \ldots, N.$$

We have

$$\mathcal{L}(\zeta_j, \tau\sqrt{p}) \le \mathcal{L}(\sum_{k\in\sigma(\mathbf{x})} x_k\xi_{jk}X_{jk}, \tau\sqrt{p}).$$

Using a Berry-Esseen bound we get

$$\mathcal{L}(\zeta_j, \tau\sqrt{p}) \le C\tau + C\frac{\sum_{k\in\sigma(x)} x_k^3 p\,\mathbb{E}\,|X_{jk}|^3}{(\sum_{k\in\sigma(\mathbf{x})} x_k^2 p)^{\frac{3}{2}}} \le C\tau + \frac{C\mu_3}{\rho\sqrt{n\delta_{0N}p}}.$$

Note that $np\delta_{0N} = y\delta_0 Np^2/(1 + |\ln p|)$. Choosing D sufficiently large, we have

$$\mathcal{L}(\zeta_j, \tau\sqrt{p}) \le 1 - b,$$

for some constant $b \in (0, 1)$. By Lemma 3.8 we get

$$\Pr\{\|\mathbf{X}\mathbf{x}\|_2 \le 2\tau\sqrt{Np}\} \le \exp\{-cN\},$$

for $\tau \le \tau_0$ and $c > 0$.

Inequality (3.2) implies that there exists $\gamma_0 > 0$ such that

$$\Pr\{\inf_{\mathbf{x}\in\mathcal{C}_1\cap \mathit{Incomp}(\delta_0,\rho)} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\} \le \exp\{-cN\}.$$

Note that

$$\mathit{Comp}(\delta_{LN}, \rho) \subset \mathcal{C}_0 \cup \left(\cup_{\nu=1}^L \widehat{\mathcal{C}}_\nu\right).$$

Using a union bound, we get

$$\Pr\{\inf_{x\in \mathit{Comp}(\delta_{LN},\rho)} \|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{np}\} \le \exp\{-cNp\} + \sum_{\nu=1}^{L-1} \exp\{-c(Np)^\nu N\delta_{0,N}\} \le \exp\{-\overline{c}Np\}. \tag{3.38}$$

By Corollary 3.3,

$$\mathrm{Comp}(\gamma_0, \rho) \subset \mathcal{C}_L.$$

This implies that

$$\inf_{\mathbf{x}\in Incomp(\gamma_0,\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le \inf_{\mathbf{x}\in Incomp(\delta_{LN},\rho)}\|\mathbf{X}\mathbf{x}\|_2. \tag{3.39}$$

In what follows we shall estimate the probability $\Pr\{\inf_{\mathbf{x}\in Incomp(\gamma_0,\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np}\}$.

3.3. **Incompressible Vectors.** Using a decomposition of the unit sphere $\mathbb{S}^{(n-1)} = Comp \cup Incomp$, we decompose the invertibility problem onto two sub problems for compressible and incompressible vectors:

$$\begin{aligned}&\Pr\{s_n(\mathbf{X}) \le \varepsilon\sqrt{p}\sqrt{N}\}\\ &\qquad\le \Pr\{\inf_{\mathbf{x}\in Comp}\|\mathbf{X}\mathbf{x}\|_2 \le \varepsilon\sqrt{p}\sqrt{N}\} + \Pr\{\inf_{\mathbf{x}\in Incomp}\|\mathbf{X}\mathbf{x}\|_2 \le \varepsilon\sqrt{p}\sqrt{N}\}.\end{aligned} \tag{3.40}$$

A bound for the compressible vectors follows from inequality (3.38). It remains to find a lower bound for $\|\mathbf{X}\mathbf{x}\|_2$ for incompressible vectors. Let $\eta, \eta_1, \ldots, \eta_N$ denote standard Gaussian random variables independent of X_{jk}, ξ_{jk} for $1 \le j \le N, 1 \le k \le n$. We shall prove the following lemma.

Lemma 3.14. *Let $x \in \mathcal{IC}(\delta,\rho)$. Then there exist absolute constants c_1 such that for any $C > 0$, the following inequality*

$$\Pr\{\|\mathbf{X}x\|_2 \le t\sqrt{Np}\} \le \Big(\frac{2t}{\sqrt{t^2+\rho^2/2}}\Big)^N + \Big(\frac{2c_0}{C}\exp\{-\frac{C^2}{2}\}\Big)^N, \tag{3.41}$$

holds for $t \ge c_1\mu_4/\sqrt{Np\delta}$.

Proof. We may write

$$\Pr\{\|\mathbf{X}x\|_2 \le t\sqrt{Np}\} = \Pr\{\sum_{j=1}^N \zeta_j^2 < t^2Np\} \tag{3.42}$$

where $\zeta_j = \sum_{k=1}^n X_{jk}\xi_{jk}x_k$. Applying Markov's inequality, we get

$$\Pr\{\sum_{j=1}^N \zeta_j^2 < t^2Np\} \le \mathrm{e}^N\,\mathbb{E}\exp\{-\frac{1}{t^2p}\sum_{j=1}^N\zeta_j^2\} = \mathrm{e}^N\prod_{j=1}^N\mathbb{E}\exp\{-\frac{1}{t^2p}\zeta_j^2\}. \tag{3.43}$$

We may rewrite the r.h.s. of (3.43) as follows

$$\Pr\{\sum_{j=1}^N \zeta_j^2 < t^2Np\} \le \mathrm{e}^N\prod_{j=1}^N\mathbb{E}\exp\{i\frac{1}{t\sqrt{p}}\zeta_j\eta_j\}. \tag{3.44}$$

Conditioning by η_j, we get

$$\Pr\{\sum_{j=1}^N \zeta_j^2 < t^2Np\} \le \mathrm{e}^N\prod_{j=1}^N\mathbb{E}_{\eta_j}\prod_{k=1}^n|\,\mathbb{E}_{X_{jk}\xi_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\xi_{jk}\}| \tag{3.45}$$

By Lemma 3.11 there exists a set $\sigma(x)$ such that for $k\in\sigma(x)$ we have $\frac{1}{2\sqrt{n}}\le|x_k|\le\frac{\sqrt{2}}{\sqrt{n\delta}}$ and $|\sigma(x)|\ge\frac{1}{2y}\delta\rho^2N$. We may write the following inequality

$$\mathbb{E}_{\eta_j}\prod_{k\in\sigma(x)}|\,\mathbb{E}_{X_{jk}\xi_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\xi_{jk}\}|\\ \le\mathbb{E}_{\eta_j}\prod_{k\in\sigma(x)}|\,\mathbb{E}_{X_{jk}\xi_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\xi_{jk}\}|. \tag{3.46}$$

For any constant C we have

$$\mathbb{E}_{\eta_j}\prod_{k\in\sigma(x)}|\,\mathbb{E}_{X_{jk}\xi_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\xi_{jk}\}|\\ \le\mathbb{E}_{\eta_j}\left(\prod_{k\in\sigma(x)}|\,\mathbb{E}_{X_{jk}\xi_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\xi_{jk}\}|\right)\mathbb{I}\{|\eta_j|\le C\}+\Pr\{|\eta_j|>C\}. \tag{3.47}$$

Consider $k\in\sigma(x)$ now. Taking expectation with respect to ξ_{jk} conditioning on X_{jk} and η_j), we obtain

$$|\,\mathbb{E}_{X_{jk}\xi_{jk}}\big(\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\xi_{jk}\}\big)|\\ =|1+p(\mathbb{E}_{X_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\}-1)|. \tag{3.48}$$

Applying Taylor's formula for the characteristic function $\mathbb{E}_{X_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\}$, we may write

$$|1+p(\mathbb{E}_{X_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}||\eta_j|\le C\}-1)|\\ \le|1+p(-\frac{1}{2t^2p}\eta_j^2x_k^2+\frac{\mathbb{E}\,|X_{11}|^3}{6t^3p^{\frac32}}|x_k|^3|\eta_j|^3)|. \tag{3.49}$$

Since $\mathbb{E}\,|X_{11}|^3\le\mathbb{E}^{\frac34}\,|X_{11}|^4\le\mu_4^{\frac34}\le\mu_4$, for $|\eta_j|\le C$, and

$$t\ge\frac{C\mu_4}{\sqrt{yNp\delta}}, \tag{3.50}$$

we have

$$\frac{|x_k||\eta_j|\,\mathbb{E}\,|X_{11}|^3}{3t\sqrt{p}}\le\frac{C\mu_4\sqrt{2}}{3t\sqrt{yN\delta p}}\le\frac12.$$

Taking into account this inequality, we get for $|\eta_j|\le C$,

$$|1+p(\mathbb{E}_{X_{jk}\xi_{jk}}\exp\{i\frac{1}{t\sqrt{p}}\eta_jx_kX_{jk}\}-1)|\le\exp\{-\frac{1}{4t^2}x_k^2\eta_j^2\}. \tag{3.51}$$

Since $\sum_{k\in\sigma(x)} x_k^2 \ge \rho^2$, this inequality implies that

$$\prod_{k=1}^{n} |\,\mathbb{E}_{X_{jk}\xi_{jk}} \exp\{i\frac{1}{t\sqrt{p}}\eta_j x_k X_{jk}\xi_{jk}\}|\mathbb{I}\{|\eta_j|\le C\} \le \exp\{-\frac{\rho^2}{4t^2}\eta_j^2\}. \tag{3.52}$$

From here it follows for any $C>0$

$$\Pr\{\sum_{j=1}^{N}\zeta_j^2 < t^2Np\} \le \prod_{j=1}^{N}\big(\mathbb{E}\exp\{-\frac{\rho^2}{4t^2}\eta_j^2\} + \Pr\{|\eta_j|>C\}\big). \tag{3.53}$$

There exists an absolute constant $c_0>0$ such that

$$\Pr\{|\eta_j|>C\} \le \frac{c_0}{C}\exp\{-\frac{C^2}{2}\}. \tag{3.54}$$

This inequality implies that

$$\begin{aligned}\Pr\{\sum_{j=1}^{N}\zeta_j^2 < t^2Np\} &\le (\frac{t}{\sqrt{t^2+\rho^2/2}} + \frac{c_0}{C}\exp\{-\frac{C^2}{2}\})^N\\ &\le (\frac{2t}{\sqrt{t^2+\rho^2/2}})^N + (\frac{2c_0}{C}\exp\{-\frac{C^2}{2}\})^N. \end{aligned}\tag{3.55}$$

Thus, Lemma 3.14 is proved. $\square$

Proof of Theorem 1.2:

First we note that

$$\begin{aligned}\Pr\{\inf_{\mathbf{x}\in\mathcal{S}^{(n-1)}}\|\mathbf{X}\mathbf{x}\|_2 \le t\sqrt{Np}\} &\le \Pr\{\inf_{\mathbf{x}\in\mathrm{Com}(\delta_{L,N},\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le t\sqrt{Np}\}\\ &\quad + \Pr\{\inf_{\mathbf{x}\in\mathrm{Incomp}(\delta_{L,N},\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le t\sqrt{Np}\}.\end{aligned}\tag{3.56}$$

By inequality (3.38), for some constant $\overline{c}>0$,

$$\Pr\{\inf_{\mathbf{x}\in\mathrm{Com}(\delta_{L,N},\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le t\sqrt{Np}\} \le \exp\{-\overline{c}Np\}. \tag{3.57}$$

By Relation (3.39), we have

$$\Pr\{\inf_{\mathbf{x}\in\mathrm{Incomp}(\delta_{L,N},\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le t\sqrt{Np}\} \le \Pr\{\inf_{\mathbf{x}\in\mathrm{Incomp}(\gamma_0,\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le t\sqrt{Np}\} \tag{3.58}$$

We consider an ε-net $\mathcal{N}$ on the set of incompressible vectors $\mathcal{IC}(\gamma_0,\rho)$ with $\varepsilon=\frac{t}{2K}$ where $K>0$ is fixed. It is straightforward to check that

$$\Pr\{\inf_{x\in\mathcal{IC}(\gamma_0,\rho)}\|\mathbf{X}\mathbf{x}\|_2 \le \tau\sqrt{Np},\, \|\mathbf{X}\|\le K\sqrt{Np}\} \le \Pr\{\inf_{x\in\mathcal{N}}\|\mathbf{X}\mathbf{x}\|_2 \le 2\tau\sqrt{Np}\} \tag{3.59}$$

Applying a union-bound, we get

$$\Pr\{\inf_{x\in\mathcal{N}}\|\mathbf{X}\mathbf{x}\|_2 \le 2\tau\sqrt{Np}\} \le |\mathcal{N}|\sup_{x\in\mathcal{IC}(\gamma_0,\rho)}\Pr\{\|\mathbf{X}\mathbf{x}\|_2 \le 2\tau\sqrt{Np}\}. \tag{3.60}$$

By [10, Proposition2.1], we have

$$|\mathcal{N}| \le n\left(1+\frac{2}{\varepsilon}\right)^{n-1}.$$

Then, applying the result of Lemma 3.14, we get (for $t \ge ...\frac{c_1\mu_4}{\sqrt{N\gamma_0 p}}$)

$$\begin{aligned}\Pr\{\inf_{x\in\mathcal{IC}(\gamma_0,\rho)}\|\mathbf{X}x\|_2 \le t\sqrt{Np}\} &\le |\mathcal{N}|\left((\frac{2t}{\sqrt{t^2+\rho^2/2}})^N + (\frac{2c_0}{C}\exp\{-\frac{C^2}{2}\})^N\right)\\ &\le yN\left(1+\frac{4K}{t}\right)^{n-1}\left((\frac{2t}{\sqrt{t^2+\rho^2/2}})^N + (\frac{2c_0}{C}\exp\{-\frac{C^2}{2}\})^N\right). \end{aligned} \tag{3.61}$$

It is easy to see that, for any $0 < t \le \tau_0$,

$$\Pr\{\inf_{x\in\mathcal{IC}(\delta,\rho)}\|\mathbf{X}x\|_2 \le t\sqrt{Np}\} \le \Pr\{\inf_{x\in\mathcal{IC}(\delta,\rho)}\|\mathbf{X}x\|_2 \le \tau_0\sqrt{Np}\}. \tag{3.62}$$

Without loss of generality we may assume that $\tau_0 \le 4K$. Taking into account both that $N \le \mathrm{e}^N$ and $y < 1$ rewrite the inequality (3.63) in the form

$$\begin{aligned}\Pr\{\inf_{x\in\mathcal{IC}(\gamma_0,\rho)}\|\mathbf{X}x\|_2 \le \tau_0\sqrt{Np}\} &\le \left(\frac{5K}{2\tau_0}\right)^{yN}\left((\frac{4\mathrm{e}\tau_0}{\sqrt{4\tau_0^2+\rho^2/2}})^N + (\frac{2c_0\mathrm{e}}{C}\exp\{-\frac{C^2}{2}\})^N\right)\\ &\le \left(\left(\frac{(5K)^y4\mathrm{e}\sqrt{2}}{\rho}\tau_0^{(1-y)}\right)^N + (\frac{2c_0\mathrm{e}(5K)^y}{C\tau_0^y}\exp\{-\frac{C^2}{2}\})^N\right)\end{aligned} \tag{3.63}$$

Put

$$\tau_0 = \left(\frac{\rho}{4\sqrt{2}\cdot 5^y\mathrm{e}^2K^y}\right)^{\frac{1}{1-y}}.$$

For $N \ge 2$, we have

$$\frac{(5K)^y4\mathrm{e}\sqrt{2}}{\rho}\tau_0^{(1-y)} \le \frac{1}{2}\mathrm{e}^{-\frac{1}{2}N}. \tag{3.64}$$

Note that, by condition (1.5), for N such that

$$\ln N \ge \frac{\mu_4}{\tau_0\sqrt{B\gamma_0}}, \tag{3.65}$$

we have

$$\tau_0 \ge \frac{\mu_4}{\sqrt{Np\gamma_0}}. \tag{3.66}$$

Moreover, choosing C such that

$$C\mathrm{e}^{\frac{C^2}{2}} \ge \frac{2c_0\mathrm{e}5^yK^y}{2^y\rho^{\frac{y}{1-y}}\tau_0^y}$$

we obtain that

$$\Pr\{\inf_{x\in\mathcal{IC}(\gamma_0,\rho)}\|\mathbf{X}x\|_2 \le t\sqrt{Np}, \|\mathbf{X}\| \le K\sqrt{Np}\} \le \mathrm{e}^{-N/2}, \tag{3.67}$$

for any $0 \le t \le \tau_0$. The result of Theorem 1.2 follows now from inequalities (3.56), (3.57) and (3.67). (Since γ_0 is an absolute constant defined in Corollary 3.3 .) Theorem 1.2 is proved.

References

[1] Roland Bauerschmidt, Jiaoyang Huang, Antti Knowles, Horng-Tzer Yau. *Edge rigidity and universality of random regular graphs of intermediate degree.* Geometric and Functional Analysis 30, 693–769, (2020), DOI: 10.1007/s00039-020-00538-0.

[2] Paul Bourgade, Jiaoyang Huang, Horng-Tzer Yau. *Eigenvector statistics of sparse random matrices.* Electron. J. Probab. 22: 1–38 (2017). DOI: 10.1214/17-EJP81, eprint arXiv:1609.09022

[3] Xiucai Ding, Fan Yang. *A necessary and sufficient condition for edge universality at the largest singular values of covariance matrices.* Ann. Appl. Probab., 28(3): 1679–1738, 2018.

[4] Laszlo Erdös, Antti Knowles, Horng-Tzer Yau, Jun Yin. *Spectral Statistics of Erdös – Renyi Graphs I: Local Semicircular Law.* The Annals of Probability, 2013, Vol. 41, No. 3B, 2279–2375 DOI: 10.1214/11-AOP734.

[5] Friedrich Götze, Alexander N. Tikhomirov. *On the circular law.* Annals of Probability, 2010, vol. 38, 1444–1491.

[6] Friederich Götze, Alexey A. Naumov, Alexander N. Tikhomirov. *Moment inequalities for linear and nonlinear statistics.* Teor. Veroyatnost. i Primenen., v.65, issue 1, p. 3–22, 2020.

[7] Jong Yun Hwang anf Ji Oon Lee. *Local Law and Tracy – Widom Limit for Sparse Sample Covariance Matrices.* Bernoulli 26(3):2400–2435 (2020).

[8] Alexander Litvak, Alain Pajor, Mark Rudelson, Nicole Tomczak-Jaegermann. *Smallest singular value of random matrices and geometry of random polytopes.*Adv. Math. 195,(2005), 491–523.

[9] Alexander E. Litvak, Konstantin E. Tikhomirov. *Singularity of sparse Bernoulli matrices*, Duke Mathematical Journal, 2022, 1135–1233 (1 April 2022). DOI: 10.1215/00127094-2021-0056.

[10] Mark Rudelson, Roman Vershynin. *The smallest singular value of a random rectangular matrix.* Communications on Pure and Applied Mathematics 62(12), 2009, pp.1707–1739, DOI: 10.1002/cpa.20294.

[11] Mark Rudelson, Roman Vershynin. *The Littlewood -Offord problem and invertibility of random matrices.* Advances in Mathematics, 218(2008), 600–633.

[12] Mark Rudelson, Konstantin Tikhomirov. *Sparse circular law under minimal assumptions.* Geometric and Functional Analysis, 29, pages 561–637(2019).

[13] Yoav Seginer. *The expected norm of random matrices.* Combinatorics. Probability and Computing (2000) 9, 149–166.

[14] Konstantin E. Tikhomirov. *Singularity of random Bernoulli matrices* Annals of mathematics, vol. 191, pp 592–639.

[15] Konstantin E. Tikhomirov. *The smallest singular value of a random rectangular matrix with no moment assumptions on entries.* Israel Journal of Mathematics volume 212, pages 289–314 (2016).

Friedrich Götze, Faculty of Mathematics, Bielefeld University, Bielefeld, Germany

Email address: goetze@math.uni-bielefeld.de

Alexander N. Tikhomirov, Institute of Physics and Mathematics, Komi Science Center of Ural Branch of RAS, Syktyvkar, Russia;

Email address: tikhomirov@ipm.komisc.ru