Singularity of Bernoulli matrices in the sparse regime

\[ pn = O(\log(n)) \]

Han Huang

September 30, 2020

Abstract

Consider an \( n \times n \) random matrix \( A_n \) with i.i.d Bernoulli(\( p \)) entries. In a recent result of Litvak-Tikhomirov, they proved the conjecture

\[ \mathbb{P}\{ A_n \text{ is singular} \} = (1 + o_n(1))\mathbb{P}\{ \text{either a row or a column of } A_n \text{ equals zero} \}. \]

for \( C \frac{\log(n)}{n} \leq p \leq \frac{1}{C} \) for some large constant \( C > 1 \). In this paper, we setted this conjecture in the sparse regime when \( p \) satisfies

\[ 1 \leq \liminf_{n \to \infty} \frac{pn}{\log(n)} \leq \limsup_{n \to \infty} \frac{pn}{\log(n)} < +\infty. \]

1 Introduction

The question about the singularity of discrete random matrices has been studied for more than half of a century. One of the central problems is the probability estimate of the singularity of \( n \times n \) random matrix with i.i.d entries \( \pm 1 \). This question was firstly tackled by Komlós [4], who shows that the probability decays to 0 as \( n \) tends to infinity in 1967. For the value of the probability, it was conjectured that the probability should be \( (\frac{1}{2} + o_n(1))^n \). Then, in 1995, Kahn-Komlós-Szeméredi [3] shows that the probability is bounded above by \( \exp(-cn) \). Later it was improved by Tao-Vu [16, 17] and Bourgain-Vu-Wood [2]. In the end, this conjecture was settled by a recent result of Tikhomirov [19] in 2018:

**Theorem 1.1.** Let \( B_n \) be the \( n \times n \) random matrices with i.i.d. \( \pm 1 \) entries. For any \( \epsilon > 0 \), there exists \( C > 0 \) such that

\[ \mathbb{P}\left\{ s_{\min}(B_n) \leq tn^{-1/2} \right\} \leq Ct + C\left(\frac{1}{2} + \epsilon\right)^n \quad t \geq 0. \]

What happens when the entries are no longer balanced? There is a conjecture for Bernoulli matrices:

**Conjecture 1.2** (Stronger singularity conjecture for Bernoulli matrices). Let \( A_n \) be a \( n \times n \) random matrices with i.i.d Bernoulli(\( p \)) entries where \( p := p_n \in (0, \frac{1}{2}] \). Then,

\[ \mathbb{P}\{ A_n \text{ is singular} \} = (1 + o_n(1))\mathbb{P}\{ \text{a row or a column of } A_n \text{ equals zero, or two rows or columns are equal} \}. \]

In the regime that \( \limsup p_n < 1/2 \) then

\[ \mathbb{P}\{ A_n \text{ is singular} \} = (1 + o_n(1))\mathbb{P}\{ \text{either a row or a column of } A_n \text{ equals zero} \}. \]
This conjecture is partially resolved by a paper of Litvak-Tikhomirov [20]:

**Theorem 1.3.** There is a universal constant $C > 1$ with the following property. Let $A_n$ be an $n \times n$ random matrix whose entries are i.i.d Bernoulli($p$), with $p = p_n$ satisfying

$$C \log(n) \leq np \leq C^{-1}.$$

Then, when $n$ is sufficiently large,

$$\mathbb{P}\{A_n \text{ is singular}\} = (1 + o_n(1))\mathbb{P}\{A_n \text{ contains a zero row of column}\} = (2 + o_n(1))n(1-p)^n,$$

where $o_n(1)$ is a term which vanishes as $n \to \infty$.

Quantitatively, for $t > 0$,

$$\mathbb{P}\{s_{\min}(A) \leq t \exp(-3 \log^2(2n))\} = t + (2 + o_n(1))n(1-p)^n.$$ 

Further, this estimate can be improved when $p$ is also bounded below by a constant. Let $q \in (0, C^{-1})$ be a parameter. Then, there exists $C_q > 0$ such that if $p \geq q$,

$$\mathbb{P}\{s_{\min}(A) \leq C_q n^{-2.5} t\} = t + (2 + o_n(1))n(1-p)^n.$$ 

for sufficiently large $n$ which may depends on $q$.

In other words, the theorem settles the Conjecture 1.2 for majorities of $p$, but the two extreme cases are still open: The first case is when $c \leq p \leq 1/2$ for a small constant $c > 0$. The second case is when $p$ is of order $O(\log(n)/n)$ and close to $\log(n)/n$. Notice that in the case $p < \log(n)/n$, Conjecture 1.2 becomes trivial because the chance of having a 0 column or row is a high probability event.

Further, an earlier result of Basak-Rudelson [1] on singularity of sparse random matrices implies that the Conjecture 1.2 is true when $\log(n) \leq p_n n \leq \log(n) + o(\log(\log(n)))$. So there is a remaining gap in the following regime:

$$1 \leq \lim \inf_{n \to \infty} \frac{pn}{\log(n)} \leq \lim \sup_{n \to \infty} \frac{pn}{\log(n)} < +\infty.$$ 

In this paper, we show Conjecture 1.2 is true in the above regime.

The following notations $n, p$ and $A$ will be consistent in this paper:

Let $A$ be a $n \times n$ matrices with i.i.d Bernoulli entries with probability $p$ where $p = p_n$ is a $n$-dependent value satisfies

$$1 \leq \lim \inf_{n \to \infty} \frac{pn}{\log(n)} \leq \lim \sup_{n \to \infty} \frac{pn}{\log(n)} < +\infty. \quad (1)$$

**Theorem 1.4.**

$$\mathbb{P}\{A \text{ is singular}\} = (1 + o_n(1))\mathbb{P}\{A \text{ contains a zero row of column}\} = (1 + o_n(1))\left(1 - (1 - (1 - p)^n)^{2n}\right).$$ 

Moreover, for $t > 0$,

$$\mathbb{P}\{s_{\min}(A) \leq tn^{-2+o_n(1)}\} = t + (1 + o_n(1))\left(1 - (1 - (1 - p)^n)^{2n}\right).$$
We remark that \((1 - (1 - (1 - p)^n))^{2n} = (1 + o_n(1))2n(1 - p)^n\) when \(\lim \inf_{n \to \infty} \frac{pn}{\log(n)}\) is strictly greater than 1, but no longer holds when \(pn\) is sufficiently close to \(\log(n)\). Instead, the following is true: \((1 - (1 - (1 - p)^n))^{2n} = O_n(n(1 - p)^n)\).

Further, the interpretation of \(1 - (1 - (1 - p)^n))^{2n}\) is simple: \((1 - (1 - p)^n))^{2n}\) will be the probability that every column and row of \(A\) is not \(0\), if we falsely assume that the columns and rows are jointly independent.

And the equality

\[
P\{A \text{ contains a zero row of column }\} = (1 + o_n(1)) (1 - (1 - (1 - p)^n))^{2n}\]

becomes less straightforward when \(pn\) is close to \(\log(n)\). It will be presented in Lemma 2.5.

In the framework of Litvak-Tikhomirov [20], the authors break vectors in \(\mathbb{R}^n\) into several groups. And for each group, the authors estimate the norm of \(A_nx\) for \(x\) in that group with different approaches involving Littlewood-Offord type theorems and Geometric tools. This strategy could be traced back to the work of Tao-Vu [18] and Rudelson [14], and it has been a line of research that tackles the least singular value estimate very successfully. The optimal tail bound for general random matrices has been obtained by Rudelson-Vershynin [15]. They show that

\[
P\{s_{\min}(M_n) \leq t\} \leq C t + \exp(-cn)
\]

for an \(n \times n\) random matrix \(M_n\) with i.i.d, mean 0, unit variance, subgaussian entries. The later papers of of Rebrova-Tikhomirov [13], Livishyt [11], and Livishyt-Tikhomirov-Vershynin [12] have removed these constraints on entries. In the regime where the entries are no longer independent, there is a sequence of works \([6, 7, 5, 8, 10, 9]\) by the authors Litvak, Lytova, Tikhomirov, Tomczak-Jaegermann, and Youssef on adjacency matrices of random regular graphs. In Litvak-Tikhomirov [20], its decomposition of \(\mathbb{R}^n\) is based on this sequence of work on dealing with adjacency matrices. In particular, \(\mathbb{R}^n\) has been decomposed into three types of vectors: Gradual non-constant vectors, \(\mathfrak{A}\)-vectors, and steep vectors. We shall briefly describe how they manage each type of vectors and leave the detail to the next section.

For \(x \in \mathbb{R}^n\), let \(\sigma_x: [n] \mapsto [n]\) be a permutation such that \(|x_{\sigma_x(i)}| \geq |x_{\sigma_x(j)}|\) for \(1 \leq i \leq j \leq n\). And let \(x^* \in \mathbb{R}^n\) be the non-increasing rearrangement of \(x\) defined in the following way: \(x_i^* = |x_{\sigma_x(i)}|\).

Without explicit definition, a gradual vector \(x\) is a vector that the growth of \(x_i^*\) as \(i\) tends to 1 is relatively stable. A non-constant vector is a vector whose components are not approximately the same.

Gradual non-constant vectors represent the majority of vectors in \(\mathbb{R}^n\). The way to treat gradual non-constant vectors is based on "Littlewood-Offord" type of results. The "Littlewood-Offord" type theorems study the Lévý Concentration of the inner product of a random vector \(\xi\) with a fixed vector \(v\). For a random variable \(X\) and \(t > 0\), the Lévy Concentration of \(X\) with parameter \(t\) is defined by

\[
\mathcal{D}(X, t) = \max_{s \in \mathbb{R}} P\{|\xi \cdot v - s| \leq t\}.
\]

The value \(\mathcal{D}(\xi \cdot v, t)\) depends not only on \(\xi\), but also on the structure of \(v\). For example if \(\xi\) is a random vector with i.i.d \(\pm 1\) components, the probability that \(\xi \cdot v = 0\) has a significant difference between the case when \(v = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \ldots, 0)\) and the case when \(v = (\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}})\), while \(v\) has norm 1 in both cases. In the estimate of least singular value, the Lévy concentration was used for estimating the lower bound of \(|C_i(M_n) \cdot Y|\) for \(i \in [n]\), where \(C_i(M_n)\) is the \(i\)th column of \(M_n\) and \(Y\) is a unit normal vector of the subspace spanned by \(C_j(M_n)\) with \(j \neq i\). (Notice that, typically, if \(C_i(M_n) \cdot Y = 0\), then the matrix is not invertible.)

In the paper of Litvak-Tikhomirov [20], they developed a new Littlewood-Offord type result, which focuses on Bernoulli matrices and provides a better probability estimate which works for \(p\) from constant regime down to the level \(p = O(1/n)\). Roughly speaking, this treatment leads to the conclusion that with high probability, \(\|A_n x\|\) has a proper lower bound for every gradual non-constant vector \(x \in \mathbb{R}^n\).
\( \mathcal{R} \)-vectors is the collection of vectors which one can apply technique involving 2 component: Individual probability estimate on the lower bound of \( \| A_n x \| \) via Rogozin’s Theorem 2.2 and a geometric net argument. For a fixed vector \( x \), Rogozin’s Theorem allows us to estimate the magnitude of the component \( (A_n x)_i = \sum_{j \in [n]} a_{ij} x_j \). To see that, let us take the \( n \times n \) Bernoulli(\( p \)) matrix \( A_n = (a_{ij})_{i,j \in [n]} \) as the example. First, we know that \( \mathcal{P}(a_{ij} \leq 1 - p \text{ for any } t \in (0, 1) \text{ and } j \in [n]) \). Suppose \( x \in \mathbb{R}^n \) is a vector such that
there exists \( J \in [n] \) with small value of \( \| x_J \|_\infty / \| x_J \|_2 \), where \( x_J \) is the obtained by restricting the indices of \( x \) to \( J \). Rogozin’s Theorem 2.2 implies
\[
\mathbb{P} \{ |(A x)_i| \leq \| x_J \|_\infty t \} \leq \frac{C}{\sqrt{p}} \| x_J \|_2.
\]
Then, together with a tensorization trick, one is able to show \( \| A x \| \) is not small with large probability. Examing the inequality (2), we know it is necessary that there exists \( J \subset [n] \) with \( \| x_J \|_\infty / \| x_J \|_2 < c \sqrt{p} \) for the inequality to be useful. And thus it give a restriction on what type of vectors can be a \( \mathcal{R} \)-vectors. On the other hand, the collection of \( \mathcal{R} \)-vectors need to have low complexity: If we properly normalized all \( \mathcal{R} \)-vectors, it should be able covered by a net with small cardinality (and that is why we cannot include majority of vectors in \( \mathbb{R}^n \) as \( \mathcal{R} \)-vectors).

The last component is steep vectors, or \( T \)-vectors. One can view it as the counterpart of gradual vectors. For instance, it can be characterized by the following property:
\[
x_k^* > C_{p,n} x^m
\]
for various choices of \( 1 \leq k \leq m \leq n \) and large \( C > 1 \). The way we try to show the norm of \( A_n x \) is away from 0 is essentially showing the existence of a row \( R_i(A_n) \) with the following property: There exists \( j_0 \in \{ \sigma_x(l) \}_{l \in [k]} \) such that \( a_{ij_0} = 1 \) and \( a_{ij} = 0 \) for \( j \in \{ \sigma_x(l) \}_{l \in [m]} \setminus \{ j_0 \} \). If \( C_{p,n} \) is large enough, then \((A x)_i\) is dominated by the term \( a_{ij_0} x_{j_0}^* \), whose magnitude can be bounded below via \( x_k^* \).

The main difficulty in Livtak-Tikhomirov [20] to extend their result to the case when \( p \) gets close to \( \log(n)/n \) comes from the probability estimate for dealing with steep vectors.

By (2), one could tell that Rogozin’s Theorem cannot handle the collection of sparse vectors with support size less than \( O(1/\sqrt{p}) \). These vectors will fall into the collection of steep vectors. In the work of Basak-Rudelson [1], which study singularity of sparse random matrices, while their decomposition is different, they also faced the similar issue when dealing with sparse vectors. The treatment they have relies on the expansion property, which turns out that it is the columns of the matrices with small support size that causes the problem on the probability estimate.

To illustrate what we mean in the above sentence. Let us consider a simple example. Suppose we want to show that with high probability, for every \( J = (j_1, j_2) \subset [n] \), there exists \( i \in [n] \) such that such that \( (a_{ij_1}, a_{ij_2}) = (1, 0) \) or \( (0, 1) \). This will take care of those vectors satisfies (3) with \( k = 2 \) and \( m = 3 \). A simple calculation shows that for a fixed \( J = (j_1, j_2) \), the probability of no such \( i \) exists is about \( \exp(-2pn + O(p^2n)) \). On the other hand, suppose the corresponding columns \( C_{j_1}(A_n) \) and \( C_{j_2}(A_n) \) to have typical support size, say between \( \frac{1}{2}pn \) and \( 2pn \). The condition that no such \( i \) exists is equivalent to the supports of these two columns are not the same. In particular, they need to have the same size if there is no such \( i \). Now we assume that both columns has support size \( cpn \). The probability that their supports are the same is \( \frac{n}{c pn} = \exp(-C \log(1/p) pn) \), which we gain an extra \( \log(1/p) \) in the exponent comparing to the direct estimate.

The proof of Theorem 1.4 follows the framework of Litvak-Tikhomirov [20]. We will extract the part which is related to gradual non-constant portion from Litvak-Tikhomirov [20] as a black box. Our main effort is to handle the \( \mathcal{T} \)-vectors, where the proof is developed from the idea we mentioned above, which appeared in Basak-Rudelson [1].

The remaining of this paper is structured in the following way:
Section 2: Notations, Tools, and standard Probability Estimates

For a positive integer $m$, $[m]$ denotes the set $\{1, 2, \ldots, m\}$.

For an $m_1 \times m_2$ matrix $A = (a_{ij})_{i \in [m_1], j \in [m_2]}$, let $R_i(A)$ and $C_j(A)$ be its $i$th row and $j$th column. For a subset $J \subset [m_2]$, $A_J = (a_{ij})_{i \in [m_1], j \in J}$ is the submatrix of $A$ whose columns are restricted to the index set $J$. Furthermore, we will abuse the notation by setting $A_{i_0,j} = (a_{i_0,j})_{j \in J}$ for $i_0 \in [m_1]$ and $J \subset [m_2]$. Further, for non-negative integer $k \geq 0$, let

$$L_A(k) := |\{i \in [m] : |\text{supp}(C_iA)| \leq k\}|.$$  \hspace{1cm} (4)

First, we will cite a norm estimate for sparse Bernoulli matrices from Litvak-Tikhomirov [20]:

**Lemma 2.1.** For every $s > 0$ and $R \geq 1$, there exists a constant $C_{\text{norm}} \geq 1$ depending on $s,R$ with the following property. Let $n \geq \frac{16}{s}$ be large enough and $p \in (0,1)$ satisfies $s \log(n) \leq pn$. Let $A$ be a $n \times n$ Bernoulli(p) matrix. Then,

$$\mathbb{P}\{|A - \mathbb{E}A| \geq C_{\text{norm}}\sqrt{pn} \text{ or } |A| \geq C_{\text{norm}}\sqrt{pn} + pn\} \leq \exp(-Rpn).$$

Next, the following is the Theorem of Rogozin on anticoncentration:

**Theorem 2.2.** Consider independent random variables $X_1, \ldots, X_n$ and $\lambda_1, \ldots, \lambda_n > 0$. For $\lambda > \max_{i \in [n]} \lambda_i$, we have

$$\mathcal{D}\left(\sum_{i \in [n]} X_i, \lambda\right) \leq \frac{C\lambda}{\sqrt{\sum_{i \in [n]} \lambda_i^2(1 - \mathcal{D}(X_i, \lambda_i))}}$$

where $C_{\text{Rgz}} > 0$ is an universal constant independent from $n, X_1, \ldots, X_n$ and $\lambda, \lambda_1, \ldots, \lambda_n$.

As a consequence, let us consider the following: Let $x \in \mathbb{R}^n$ and $\xi_1, \ldots, \xi_n$ be i.i.d. Bernoulli random variables with parameter $p$. For $I \subset [n]$ and $\lambda > \|x_I\|_\infty$, we have

$$\mathcal{D}\left(\sum_{i \in [n]} x_i\xi_i, \lambda\right) \leq \mathcal{D}\left(\sum_{i \in I} x_i\xi_i, \lambda\right) \leq \frac{C_{\text{Rgz}}\lambda}{\sqrt{\sum_{i \in I} x_i^2p}} \leq \frac{C_{\text{Rgz}}\lambda}{\sqrt{p}\|x_I\|}.$$  \hspace{1cm} (5)

2.1 Standard tail bound for Binomial and Hypergeometric Distributions

**Proposition 2.3.** Let $Y$ be a binomial random variable with parameter $n$ and $p \in (0, \frac{1}{2})$ (that is, $Y$ is the sum of $n$ i.i.d Bernoulli random variable with probability $p$). We have the following estimates for the tails:

$$p \leq \frac{1}{2} \quad k \geq 2pn \quad \mathbb{P}\{Y \geq k\} \leq 2\left(\frac{enp}{k}\right)^k$$  \hspace{1cm} (6)

$$p \leq \frac{1}{2} \quad k \leq \frac{1}{2}pn \quad \mathbb{P}\{Y \leq k\} \leq 2\left(\frac{enp}{k(1-p)}\right)^k(1-p)^n$$  \hspace{1cm} (7)
For simplicity, we include two special cases:

\[
p \leq \frac{1}{2} \quad \text{and} \quad \frac{pn}{5} \geq 10 \quad \Rightarrow \quad \mathbb{P} \{Y \leq \frac{pn}{5}\} \leq \exp(-\frac{1}{3}pn) \tag{8}
\]
\[
p \leq 0.01 \quad \text{and} \quad \frac{pn}{5} \geq 10 \quad \Rightarrow \quad \mathbb{P} \{Y \geq \frac{3pn}{5}\} \leq \exp(-\frac{6}{5}pn). \tag{9}
\]

Proof. Let

\[
p_k := \mathbb{P} \{Y = k\} = \binom{n}{k} p^k (1 - p)^{n-k}.
\]

The expected value of \(Y\) is \(pn\). For \(k < pn\), we will show that \(\mathbb{P} \{Y \leq k\}\) is dominated by \(p_k\). And for \(k > pn\), \(\mathbb{P} \{Y \geq k\}\) is dominated by \(p_k\).

For \(k \leq \frac{1}{2} pn\),

\[
\frac{p_{k-1}}{p_k} = \frac{k}{n-k+1} \frac{1-p}{p} \leq \frac{\frac{1}{2} pn}{n - \frac{1}{2} pn + 1} \frac{1-p}{p} \leq \frac{1}{2}.
\]

For \(k \geq 2pn\),

\[
\frac{p_{k+1}}{p_k} = \frac{n-k}{k+1} \frac{p}{1-p} \leq \frac{n - 2pn}{2pn + 1} \frac{p}{1-p} \leq \frac{1}{2}.
\]

Thus, we have \(\mathbb{P} \{Y \geq k\} \leq 2p_k\) for \(k \geq 2pn\) and \(\mathbb{P} \{Y \leq k\} \leq 2p_k\) for \(k \leq \frac{1}{2} pn\). By the standard estimate \(\binom{n}{k} \leq \left(\frac{en}{k}\right)^k\),

\[
p_k \leq \left(\frac{\exp\left(\frac{en}{k}\right)}{k(1-p)}\right)^k (1-p)^n,
\]

and \(\frac{1}{7}\) follows.

For \(\lambda \in (0, \frac{1}{2})\),

\[
p_{\lambda pn} \leq \left(\frac{e}{\lambda}\right)^{\lambda pn} \exp(-p(n - \lambda pn)) = \exp(-pnh(\lambda, p))
\]

where \(h(\lambda, p) = [1 - (1 + p + \log(\frac{1}{\lambda}))\lambda]).\) For \(p < \frac{1}{3}\) and \(\lambda = \frac{1}{3}\), \(h(\lambda, p) \geq 0.37\). we obtain \(\mathbb{P} \{Y \leq \frac{1}{2} pn\} \leq 2 \exp(-0.37pn) \leq \exp(-\frac{1}{3}pn)\) when \(pn \geq 10 \geq \frac{\log(2)}{0.1}\).

Similarly, if we take \(\lambda = 3\) and assume \(p \leq 0.001\), we have \(\mathbb{P} \{Y \geq 3pn\} \leq 2 \exp(-1.29pn) \leq \exp(-\frac{6}{5}pn)\) when \(pn \geq 10\).

\[\square\]

**Proposition 2.4.** Assume \(n\) is sufficiently large. Let \(k \leq m\) be positive integers such that \(m \leq \frac{n}{2}\) and \(k \leq \sqrt{n} \frac{2}{3}.\) Let \(I \subset [n]\) be chosen uniformly from all subsets of \([n]\) of size \(k\). Let \(Y := |I \cap [m]|\), then

\[
\mathbb{P} \{Y \geq l\} \leq C_{hg} \left(\frac{3mk}{ln}\right)^i.
\]

for a universal constant \(C_{hg} > 0\).
Proof. By a counting argument we have, for \( l \in [k] \),
\[
\mathbb{P} \{|I \cap [m]| = l\} = \binom{m}{l} \binom{n-m}{k-l} \binom{n}{k}.
\]

Recall the standard estimate for combination \( \binom{a}{b} \leq \binom{a}{b} \leq \left( \frac{ea}{b} \right)^b \) for positive integers \( a \geq b \geq 1 \). The gap \( e^k \) is too large for us. We need a slightly better bound. Recall that for \( s \geq 1 \),
\[
s! = (1 + O\left(\frac{1}{s}\right))\sqrt{2\pi} s \left(\frac{s}{e}\right)^s
\]
and
\[
\sqrt{2\pi} s \left(\frac{s}{e}\right)^s \leq s! \leq e \sqrt{s} \left(\frac{s}{e}\right)^s.
\]

By \( \binom{n}{k} = \frac{n!}{k!(n-k)!} \), we have
\[
\binom{n}{k} = (1 + o_k(1)) \frac{1}{e} \sqrt{\frac{n}{k(n-k)}} \left( \frac{n}{k} \right)^k \left( \frac{n}{n-k} \right)^{n-k}.
\]

The last term \( \left( \frac{n}{n-k} \right)^{n-k} \) can be simplified into the following form:
\[
\left( \frac{n}{n-k} \right)^{n-k} = \left( 1 + \frac{k}{n-k} \right)^{n-k} \geq e^k \exp\left( -\frac{k^2}{n-k} \right) = (1 + o_n(1)) e^k,
\]
where we used \( \log(1+x) > x - x^2 \) for \( x > 0 \) and the assumption \( \frac{k}{\sqrt{n}} = o_n(1) \). Therefore we have
\[
\binom{n}{k} \geq (1 + o_k(1)) \frac{1}{e \sqrt{k}} \left( \frac{en}{k} \right)^k.
\]

By a similar computation, for \( 1 \leq l < k \), we have
\[
\binom{n-m}{k-l} \leq \frac{1}{\sqrt{2\pi \sqrt{k-l}}} \left( \frac{e(n-m)}{k-l} \right)^{k-l},
\]
where \( \left( \frac{n-m}{n-m-k+l} \right)^{k-l} \leq e^{k-l} \) is used.

Combining these two estimates and \( \binom{m}{l} \leq \left( \frac{em}{l} \right)^l \) to get
\[
\mathbb{P} \{|I \cap [m]| = l\} \leq C \sqrt{\frac{k}{k-l+1}} \left( \frac{mk}{ln} \right)^l \left( \frac{k}{k-l} \right)^{k-l} \left( \frac{n-m}{n} \right)^{k-l} \leq C \sqrt{\frac{k}{k-l+1}} \left( \frac{emk}{ln} \right)^l,
\]
where \( C > 0 \) is some universal constant. Due to \( l \mapsto C \sqrt{\frac{k}{k-l+1}} \left( \frac{emk}{ln} \right)^l \) decays geometrically. We can find a suitable \( C_{hg} > 1 \) and replaing \( e \) by 3 to get the statement of Proposition. \( \square \)
2.2 Probability Estimate for existence of $0$ column or row

**Lemma 2.5.** Suppose $p = p_n \in (0, 1)$ satisfies $\log(n) \leq pn \leq C \log(n)$ where $C > 1$ is an arbitrary constant. Let $A$ be an $n \times n$ Bernoulli($p$) matrix. Then, when $n$ is sufficiently large,

$$\mathbb{P}\left\{ \exists i \in [n] \text{ s.t. } R_i(A) = \emptyset \text{ or } \exists j \in [n] \text{ s.t. } C_i(A) = \emptyset \right\} = (1 + o_n(1)) \left(1 - (1 - (1 - p)^n)^2n\right).$$

Furthermore, $(1 - (1 - (1 - p)^n)^2n) = O_n(n(1 - p)^n)$.

To prove Lemma 2.5, we will need the following Proposition:

**Proposition 2.6.** Let $C > 1$ be any constant. Suppose $m$ is a sufficiently large positive integer, and $p$ is a parameter satisfying $\log(m) \leq pm \leq C \log(m)$. For $0 \leq k \leq \log(m)$, let $A_k$ be a $m \times (m - k)$ Bernoulli($p$) matrix. We have

$$\mathbb{P}\{L_{A_k}(0) > 0\} = (1 + o_m(m^{-1/2}))\mathbb{P}\{L_{A_0}(0) > 0\}.$$

**Proof.** Let

$$q_k := \mathbb{P}\left\{ C_1(A) = \emptyset \right\} = (1 - p)^{m-k}.$$

From the constraints of $p$ and $k$, we have $q_k = q_0(1 + O(pk)) \leq \frac{2}{m}$.

Then,

$$\mathbb{P}\{L_{A_k}(0) > 0\} = \sum_{j=1}^{m} \binom{m}{j} q_k^j (1 - q_k)^{m-j}.$$

Our goal is to show that the summation is dominated by the first $\log(m)$ terms. For $j \geq \log(m)$,

$$\frac{\binom{m}{j+1} q_k^{j+1}(1 - q_k)^{m-j-1}}{\binom{m}{j} q_k^j(1 - q_k)^{m-j}} = \frac{m - j}{j + 1} \frac{q_k}{1 - q_k} \leq \frac{m}{j} q_k \leq 1/2$$

due to $q_k = (1 - p)^m(1 - p)^{-k} \leq \frac{2}{m}(1 + O(pk)) \leq \frac{2}{m}$. Thus, the tail of the sequence $\left\{ \binom{m}{j} q_k^j (1 - q_k)^{m-j} \right\}_{j \geq 1}$ decays geometrically at the rate $\frac{1}{2}$ when $j \geq \log(m)$.

Then, we compare the $j$th term to the 1st term with $j \geq \log(m)$:

$$\frac{\binom{m}{j} q_k^j (1 - q_k)^{m-j}}{\binom{m}{1} q_k(1 - q_k)^{m-1}} \leq \frac{1}{m} \left( \frac{em}{j} \right)^j q_k^{j-1} \leq \left( \frac{2e}{j} \right)^j \leq \frac{1}{2m},$$

where $q_k \leq \frac{2}{m}$ is applied. Therefore, we have

$$\mathbb{P}\{L_{A_k}(0) > 0\} \leq \sum_{j=1}^{\lfloor \log(m) \rfloor} \binom{m}{j} q_k^j (1 - q_k)^{m-j} + \frac{1}{m} \mathbb{P}\{L_{A_0}(0) > 0\}.$$

Next, we will compare the first $\log(m)$ summands to those in the case when $k = 0$. For $j, k \leq \log(m)$, the following estimates hold:

$$pk = o_m(m^{-1/2}), \quad q_k pk m = o_m(m^{-1/2}), \text{ and } \quad pkj = o_m(m^{-1/2}).$$
With $q_k = q_0(1-p)^{-k}$,

$$\binom{m}{j}q_k^j(1-q_k)^{m-j} = \binom{m}{j}q_0(1-q_0)^{m-j}(1+o(m^{-1/2})).$$

Therefore, we conclude that

$$\mathbb{P}\{\mathcal{L}_{Ak}(0) > 0\} = (1 + o_m(m^{-1/2}))\mathbb{P}\{\mathcal{L}_A(0) > 0\}.$$

\[\square\]

Now we are ready to prove Lemma 2.5.

**Proof of Lemma 2.5.** Let $O_R$ be the event that

$$\exists i \in [n] \text{ s.t. } \mathbf{R}_i(A) = \mathbf{0},$$

and $O_C$ be the event that

$$\exists j \in [n] \text{ s.t. } \mathbf{C}_j(A) = \mathbf{0}.$$

For $S \subset [n]$, let $P_S$ be the probability that

$$\forall i \in S, \mathbf{R}_i(A) = \mathbf{0} \text{ and } \exists j \in [n] \text{ s.t. } \mathbf{C}_j(A) = \mathbf{0}.$$

By the standard inclusion-exclusion formula, we have

$$\mathbb{P}\{O_R \cap O_C\} = \sum_{S \subset [n], S \neq \emptyset} (-1)^{|S|+1}P_S.$$

Due to the distribution of $A$ is invaraint under row permutations, $P_S$ depends only on $|S|$.

For $k \in [n]$, let $A_k$ be an $n-k$ by $k$ Bernoulli($p$) matrix. Suppose $S \subset [n]$ is a subset of size $k$. Let $O_S$ be the event that $\forall i \in S, \mathbf{R}_i(A) = \mathbf{0}$. Conditioning on $O_S$, the submatrix of $A$ obtained by restricting its rows to $[n]\setminus S$ has the same distribution as that of $A_k$. Thus, we have

$$P_S := (1-p)^{nk}\mathbb{P}\{\mathcal{L}_{Ak}(0) > 0\}$$

and we conclude that

$$\mathbb{P}\{O_R \cap O_C\} = \sum_{k \in [n]} (-1)^{k+1}\binom{n}{k}(1-p)^{nk}\mathbb{P}\{\mathcal{L}_{Ak}(0) > 0\}.$$

Notice that by the same argument we can deduce that

$$\mathbb{P}\{O_C\} = \sum_{k \in [n]} (-1)^{k+1}\binom{n}{k}(1-p)^{nk}.$$

Roughly speaking, if $\mathbb{P}\{\mathcal{L}_{Ak}(0) > 0\} = (1 + o_n(1))\mathbb{P}\{O_R\}$ for all values of $k$. Then, it implies that $\mathbb{P}\{O_R \cap O_C\} = (1 + o_n(1))\mathbb{P}\{O_R\}\mathbb{P}\{O_C\}$. In other words, the two events are approximately independent. However, it cannot be true when $k$ is large.

What we will do is to split the summation into two parts: $k < \log(n)$ and $k \geq \lceil \log(n) \rceil$.

For $k \leq \log(n)$, we will show that $\mathbb{P}\{\mathcal{L}_{Ak}(0) > 0\} = (1 + o_n(1))\mathbb{P}\{O_R\}$. This is presented separately in Proposition 2.6. And for $k > \lceil \log(n) \rceil$, we will show the summation in this part is negligible comparing
to the summation for $k < \log(n)$. Due to the signs are alternating in the summations, we need to argue in a careful way: First, show the leading term of the summation is comparable to the summation for $k < \log(n)$. Second, show that the summation for $k \geq \log(n)$ is negligible comparing to the leading term.

By Proposition 2.6,

$$\mathbb{P}\{\mathcal{L}_{A_k}(0) > 0\} = (1 + o_n(n^{-1/2}))\mathbb{P}\{O_R\}$$

for $k \leq \lceil \log(n) \rceil$. Hence,

$$\sum_{k \in [\lceil \log(n) \rceil]} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \mathbb{P}\{\mathcal{L}_{A_k}(0) > 0\}$$

$$= \sum_{k \in [\lceil \log(n) \rceil]} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \mathbb{P}\{O_R\}$$

$$+ o(n^{-1/2}) \sum_{k \in [\lceil \log(n) \rceil]} \binom{n}{k} (1-p)^{nk} \mathbb{P}\{O_R\}.$$

For $k \geq 2$,

$$\frac{\binom{n}{k}(1-p)^{nk}}{\binom{n}{k-1}(1-p)^{n(k-1)}} = n - k + 1 \cdot \frac{k}{k} (1-p)^{n} \leq \frac{1}{k} n \exp(-pn) \leq \frac{1}{k}, \quad (12)$$

where the last inequality holds due to $pn \geq \log(n)$. As an alternating sequence whose absolute values are decreasing,

$$\sum_{k \in [\lceil \log(n) \rceil]} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \mathbb{P}\{O_R\}$$

$$\geq \binom{n}{1}(1-p)^{n}\mathbb{P}\{O_R\} - \binom{n}{2}(1-p)^{2n}\mathbb{P}\{O_R\}$$

$$\geq \frac{1}{2} \binom{n}{1}(1-p)^{n}\mathbb{P}\{O_R\}$$

Notice that, (12) also implies

$$\frac{\binom{n}{k}(1-p)^{nk}}{\binom{n}{1}(1-p)^{n}} \leq \frac{1}{k!}$$

and thus

$$\sum_{k \in [\lceil \log(n) \rceil]} \binom{n}{k} (1-p)^{nk} \mathbb{P}\{O_R\} \leq e \binom{n}{1}(1-p)^{n}\mathbb{P}\{O_R\}.$$

Therefore, we conclude that

$$\sum_{k \in [\lceil \log(n) \rceil]} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \mathbb{P}\{\mathcal{L}_{A_k}(0) > 0\} \geq \frac{1}{3} \binom{n}{1}(1-p)^{n}\mathbb{P}\{O_R\}.$$. 
Now we turn to the summation for \( k \geq \lceil \log(n) \rceil \),

\[
\sum_{k \geq \lceil \log(n) \rceil} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \mathbb{P} \{ \mathcal{L}_k(0) > 0 \} 
\leq \binom{n}{1} (1-p)^n \sum_{k \in \lceil \log(n) \rceil} \frac{1}{k!} 
\leq \binom{n}{1} (1-p)^n \exp\left(-\frac{1}{2} \log(\log(n)) \log(n) \right).
\]

Next, \( \mathbb{P}(C_j(A) = 0) = (1-p)^n \) for \( j \in [n] \) and the independence of columns implies \( \mathbb{P} \{ O_R \} = 1 - (1 - (1 - p)^n)^n \).

For \( x \in (0, 1/2) \), we have

\[
1 - x = \exp(-x + O(x)^2).
\]

Conversely, for \( 0 \leq x \leq 1.5 \), \( 1 - \exp(-x) = O(x) \).

Applying these inequalities to the estimate of \( \mathbb{P} \{ O_R \} \), we obtain

\[
\mathbb{P} \{ O_R \} = 1 - \exp\left(- (1 + o_n(1)) \exp(-pn)n \right) = O(\exp(-pn)n) > \exp(-C \log(n)),
\]

(13)

where we rely on \( \exp(-pn)n \leq 1 \).

And it allows us to compare the summation for \( k \geq \lceil \log(n) \rceil \) to the first term of the summation:

\[
\sum_{k \geq \lceil \log(n) \rceil} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \mathbb{P} \{ \mathcal{L}_k(0) > 0 \} = o_n(n^{-1/2}) \left( \binom{n}{1} (1-p)^n \mathbb{P} \{ O_R \} \right).
\]

Therefore, we can conclude that

\[
\mathbb{P} \{ O_R \cap O_C \} = (1 + o_n(n^{-1/2}) \mathbb{P} \{ O_R \} . \left( \sum_{k \in \lceil \log(n) \rceil} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \right).
\]

The same approach could show that

\[
\mathbb{P} \{ O_C \} = (1 + o_n(n^{-1/2}) \left( \sum_{k \in \lceil \log(n) \rceil} (-1)^{k+1} \binom{n}{k} (1-p)^{nk} \right).
\]

Finally, we obtain

\[
\mathbb{P} \{ O_R \cap O_C \} = (1 + o_n(n^{-1/2}) \mathbb{P} \{ O_R \} \mathbb{P} \{ O_C \} = (1 + o_n(n^{-1/2})) (\mathbb{P} \{ O_R \} )^2.
\]

At this point we are ready to estimate \( \mathbb{P} \{ O_R \cup O_C \} : 
\]

\[
\mathbb{P} \{ O_R \cup O_C \} = \mathbb{P} \{ O_R \} + \mathbb{P} \{ O_C \} - \mathbb{P} \{ O_R \cap O_C \} 
= 2 \mathbb{P} \{ O_R \} - (1 + o_n(n^{-1/2})) (\mathbb{P} \{ O_R \})^2 
= (1 + o_n(n^{-1/2}) (\mathbb{P} \{ O_R \}) (2 \mathbb{P} \{ O_R \} - (\mathbb{P} \{ O_R \})^2).
\]

Substituting \( \mathbb{P} \{ O_R \} = 1 - (1 - (1 - p)^n)^n \), we obtain the statement of the lemma:

\[
\mathbb{P} \{ O_R \cup O_C \} = (1 + o(n^{-1/2} \exp(-pn)n)) (1 - (1 - (1 - p)^n)^{2n}) .
\]

It remains to show that \( \mathbb{P} \{ O_R \cup O_C \} = O_n(n \exp(-pn)) \), but it is immediate due to \( \mathbb{P} \{ O_R \} = \mathbb{P} \{ O_C \} \) and \( \mathbb{P} \{ O_R \} = O_n(n \exp(-pn)) \) from [13].

□
3 Decomposition of $\mathbb{R}^n$

Let $r \in (0, 1)$ be a parameter, and for vectors whose support is at least $|rn|$, we consider the following normalization

$$\mathcal{Y}(r) := \{ x \in \mathbb{R}^n : x^*_{[rn]} = 1 \}$$

$$\mathcal{A}(r, \rho) := \{ x \in \mathbb{R}^n : \exists \lambda \in \mathbb{R} \text{ s.t } |\lambda| = x^*_{[rn]} \text{ and } |\{ i \in [n] : |x_i - \lambda| < \rho \lambda \}| > n - |rn| \}.$$  

By a growth function $g$ we mean any non-decreasing function from $[1, \infty)$ into $[1, \infty)$. For $r > 0, \delta > 0, \rho > 0$, and a growth function $g$, we define the gradual non-constant vectors with these parameters to be the set

$$\mathcal{V}(r, g, \delta, \rho) := \left\{ x \in \mathcal{Y}(r) : \forall i \in [n], x_i^* \leq g(n/i) \text{ and } \exists Q_1, Q_2 \subset [n] \text{ such that } |Q_1|, |Q_2| \geq \delta n \text{ and } \max_{i \in Q_2} x_i \leq \min_{i \in Q_1} x_i - \rho \right\}.$$  

These parameters $r \in (0, 1), \rho > 0$, and the growth function $g$ will be chosen later.

The following is a partial result extracted from Litvak-Tikhomirov [20] which focuses only on $\mathcal{V}$-vectors:

**Theorem 3.1.** Let $r, \delta, \rho \in (0, 1), \ s > 0 \ , R \geq 1,$ and let $K_3 \geq 1$. Then, there are $n_0 \in \mathbb{N}, C \geq 1$ and $K_1 \geq 1, K_2 \geq 4$ depending on $r, \delta, \rho, R, s, K_3$ such that the following holds. Let $n \geq n_0$, $p \leq C^{-1}$, and $s \log(n) \leq pn$. Let $g : [1, \infty) \to [1, \infty)$ be an increasing growing function satisfying

$$\forall a \geq 2, \forall t \geq 1 \ g(at) \geq g(t) + a \text{ and } \prod_{j=1}^{\infty} g(2^j)^{2^j} \leq K_3.$$  

(14)

Let $A$ is an $n \times n$ Bernoulli$(p)$ random matrix and $\mathcal{E}$ be the complement of the event

$$\{ \|Ax\| \leq a_n^{-1} \|x\| \text{ or } \|A^T x\| \leq a_n^{-1} \|x\| \text{ for some } x \notin \bigcup_{\lambda \geq 0}(\lambda \mathcal{V}_n) \}$$

where $\mathcal{V}_n = \mathcal{V}(r, g, \delta, \rho)$, and $a_n > 1$. Then,

$$\mathbb{P}\left\{ \{ \|Ax\| \leq (a_n b_n)^{-1} \|x\| \text{ for some } x \in \mathcal{V}\left| \mathcal{V}_n \right) \cap \mathcal{E} \right\} \leq \exp(-2pn) + \frac{Cb_n}{r^2 a_n \sqrt{pn}} t,$$

where $b_n = \sum_{i=1}^{n} g(i)$.

3.1 $\mathcal{T}$-vectors

$\mathcal{T}$-vectors consists of three parts: $\mathcal{T}_1, \mathcal{T}_2,$ and $\mathcal{T}_3$. Recall that $p = r \frac{\log(n)}{n}$. The way we define $\mathcal{T}_1$ vectors are different when $\tau \leq 10$ or $\tau > 10$.

3.1.1 $\mathcal{T}_1$ when $1 \leq \tau \leq 10$

Let $t_0, t_1$, and $s$ be the positive integers defined by the following formulas.

- $3^{t_0-1} < \left[ \exp\left( \frac{pn}{\log^2(pn)} \right) \right] \leq 3^{t_0},$
• $\lceil \exp\left(\frac{pn}{\log^2(pn)}\right) \left(\frac{pn}{\log^3(pn)}\right)^{t_1-1} \rceil \leq \left\lceil \frac{1}{480e}\right\rceil \leq \lceil \exp\left(\frac{pn}{\log^2(pn)}\right) \left(\frac{pn}{\log^3(pn)}\right)^{t_1} \rceil$,

• $s = t_0 + t_1$.

Next, $n_0 = 1$. For $1 \leq j \leq s$, we set

\begin{align*}
  j &< t_0 & n_j &= 3^j \\
  j &= t_0 & n_0 &= \left\lceil \exp\left(\frac{pn}{\log^2(pn)}\right) \right\rceil \\
  t_0 + 1 \leq j \leq s - 1 & & n_j &= \left\lceil \exp\left(\frac{pn}{\log^2(pn)}\right) \left(\frac{pn}{\log^3(pn)}\right)^{j-t_0} \right\rceil \\
  j &= s & n_s &= \left\lceil \frac{1}{480e}\right\rceil
\end{align*}

Using $\frac{1}{480e} \leq n$, the following bounds on $t_0$, $t_1$, and $s$ hold:

\begin{align*}
  t_0 &\leq \frac{pn}{\log^2(pn)}, & t_1 &\leq (1 + o_n(1)) \frac{\log(n)}{\log(pn)}, \quad \text{and} \quad s \leq (1 + o_n(1)) \frac{\log(n)}{\log(pn)}. \quad (15)
\end{align*}

For $1 \leq j \leq s$, we set

\begin{align*}
  T_{1j} &:= \{ x \in \mathbb{R}^n : x \notin \bigcup_{i=1}^{j-1} T_{1i} \text{ and } x_{n_j}^* > \kappa pn x_j^* \}, \\
  T_1 &:= \bigcup_{j=1}^{s} T_{1j},
\end{align*}

where $\kappa > 1$ is a (large) universal constant which will be determined later.

**Proposition 3.2.** For $j \in [s]$ and $x \in T_{1j}$,

\begin{align*}
  \frac{\|x\|}{x_{n_j}^*} &\geq n^{1+o_n(1)}. \quad (16)
\end{align*}

And for $x \notin T_1$,

\begin{align*}
  \frac{\|x\|}{x_{n_j}^*} &\geq n^{1+o_n(1)}, \quad \frac{x_j^*}{x_{n_j}^*} \geq n^{1+o_n(1)}. \quad (17)
\end{align*}

**Proof.** We will prove the first statement, the proof for the second statement will be the same. Now we fix $j \in [s]$ and $x \in T_{1j}$. Due to the fact that $x \notin T_{1i}$ for $i \in [j - 1]$,

\begin{align*}
  \frac{x_{n_j-1}^*}{x_{n_j}^*} &\leq (\kappa pn)^{j-i}.
\end{align*}

Using the estimate $n_i \leq \left(\frac{pn}{\log^2(pn)}\right)^i$, we get

\begin{align*}
  \sum_{i=1}^{j} (x_i^*)^2 \leq \sum_{i=1}^{j} n_i (x_{n_i-1})^2 \\
  \leq \sum_{i=1}^{j} \left(\frac{pn}{\log^2(pn)}\right)^i (\kappa pn)^{2(j-i)} (x_{n_j-1})^2 \\
  = (\kappa pn)^{2j} (x_{n_j-1})^2 \left(\sum_{i=1}^{j} \left(\frac{pn}{\log^2(pn)}\right)^i (\kappa pn)^{-2i}\right) \\
  \leq (\kappa pn)^{2j} (x_{n_j-1})^2
\end{align*}
where the second to last inequality relies on the geometric decay of the summation \( \sum_{i=1}^{n_j} \left( \frac{pn}{\log^2(pn)} \right)^i (\kappa pn)^{-2i} \).

For the remaining terms,

\[
\sum_{i=n_j+1}^{n} (x_i^*)^2 \leq n \left( \frac{1}{\kappa pn} \right)^2 (x_{n_j-1}^*)^2.
\]

Combining these two inequalities we obtain

\[
\|x\|^2 \leq \left( (\kappa pn)^{2j} + \frac{n}{(\kappa pn)^2} \right) (x_{n_j-1}^*)^2
\]

and using \( j \leq s \leq (1 + o_n(1)) \frac{\log(n)}{\log(pn)} \),

\[
\|x\|^2 \leq n^{2+o_n(1)} (x_{n_j-1}^*)^2.
\]

\[\square\]

3.1.2 \( T_1 \) when \( 10 < \tau < +100 \)

Let \( s \) be the positive integer defined by the following inequalities:

\[
2 \left( \frac{pn}{\log^3(pn)} \right)^{s-2} < \left\lfloor \frac{1}{480e} \right\rfloor \leq 2 \left( \frac{pn}{\log^3(pn)} \right)^{s-1}.
\]

Let \( n_0 = 1, n_1 = 2, \) and

\[
2 \leq j < s \quad n_j = 2 \left( \frac{pn}{\log^3(pn)} \right)^{j-1}
\]

\[
j = s \quad n_s = \left\lfloor \frac{1}{480e} \right\rfloor
\]

Using \( \frac{1}{480e} \leq n \), we have

\[
s \leq (1 + o_n(1)) \frac{\log(n)}{\log(pn)}.
\]

For \( 1 \leq j \leq s \), we set

\[
T_{1j} := \left\{ x \in \mathbb{R}^n : x \notin \bigcup_{i=1}^{j-1} T_{1i} \text{ and } x_{n_j-1}^* > 24pn x_{n_j-1}^* \right\},
\]

\[
T_1 := \bigcup_{j=1}^{s} T_{1j}.
\]

The statement of the following Proposition is exactly the same as that of Proposition 3.2, except the definition of \( s \), \( T_{1j} \), and \( T_1 \) are different.

**Proposition 3.3.** For \( j \in [s] \) and \( x \in T_{1j} \),

\[
\frac{\|x\|}{x_{n_j-1}^*} \geq n^{1+o_n(1)}.
\]

And for \( x \notin T_1 \),

\[
\frac{\|x\|}{x_{n_s}^*} \geq n^{1+o_n(1)}, \quad \frac{x_{n_j}^*}{x_{n_s}^*} \geq n^{1+o_n(1)}.
\]

**Proof.** We omit the proof here since it is essentially the identical to the proof of Proposition 3.2. \[\square\]
3.1.3 \( \mathcal{T}_2, \mathcal{T}_3, \text{ and } \mathcal{T} \) (for \( 1 \leq \tau < +\infty \))

Now, we define

\[
\begin{align*}
n_{s+1} &= \frac{\sqrt{n}}{p} \\
n_{s+2} &= rn
\end{align*}
\]

where \( r \in (0, 1) \) is a sufficiently small constant. Next, let

\[
\begin{align*}
\mathcal{T}_2 &:= \{ x \in \mathbb{R}^n : x \notin \mathcal{T}_1 \text{ and } x_{n_s}^* > C_T \sqrt{pn} x_{n_{s+1}}^* \} \\
\mathcal{T}_3 &:= \{ x \in \mathbb{R}^n : x \notin \mathcal{T}_1 \cup \mathcal{T}_2 \text{ and } x_{n_{s+1}}^* > C_T \sqrt{pn} x_{n_{s+2}}^* \} \\
\mathcal{T} &:= \bigcup_{i=1}^3 \mathcal{T}_i.
\end{align*}
\]

where \( C_T > 1 \) is a large universal constant.

Proposition 3.4. For \( x \notin \mathcal{Y}(r) \setminus \mathcal{T} \), we have

\[
\begin{align*}
x_{n_{s+2}}^* &= 1 \\
x_{n_{s+1}}^* &\leq C_T \sqrt{pn} \\
x_{n_s}^* &\leq C_T^2 pn \\
x_{n_{s-j}}^* &\leq C_T^2 pn (\kappa pn)^j \text{ for } j \in [s] \\
\|x\|_\infty &\leq n^{1+o_n(1)} \\
\|x\| &\leq n^{1+o_n(1)}
\end{align*}
\]

Proof. This is an immediate consequence of Proposition 3.2, 3.3, and the definition of \( \mathcal{T}_2, \mathcal{T}_3, \text{ and } \mathcal{T} \).

\[\square\]

3.2 \( \mathcal{R} \) vectors

For \( n_s \leq k \leq \frac{n}{\log^2(pn)} \), let \( B = \{k, k+1, \ldots, n\} \) and

\[
\begin{align*}
\mathcal{R}_k^1 &:= \left\{ x \notin \mathcal{T} \setminus \mathcal{A}(\rho) : x_{rn}^* = 1, \frac{\|x_{\sigma_s(B)}\|}{\|x_{\sigma_s(B)}\|_\infty} \geq \frac{2C_{RGZ}}{r} \text{ and } \frac{n}{2} \leq \frac{\|x_{\sigma_s(B)}\|}{\|x_{\sigma_s(B)}\|_\infty} \leq C_T \sqrt{pn^2} x \right\} \\
\mathcal{R}_k^2 &:= \left\{ x \notin \mathcal{T} : x_{rn}^* = 1, \frac{\|x_{\sigma_s(B)}\|}{\|x_{\sigma_s(B)}\|_\infty} \geq \frac{2C_{RGZ}}{r} \text{ and } \frac{2\sqrt{n}}{r} \leq \frac{\|x_{\sigma_s(B)}\|}{\|x_{\sigma_s(B)}\|_\infty} \leq C_T^2 p^2 n^2 x \right\} \\
\mathcal{R}_k &:= \mathcal{R}_k^1 \cup \mathcal{R}_k^2,
\end{align*}
\]

where \( C_{RGZ} > 0 \) is the constant appeared in Theorem 2.2.

In the end, we set

\[
\mathcal{R} := \bigcup_{n_s \leq k \leq \frac{n}{\log^2(pn)}} \mathcal{R}_k.
\]
3.3 Gradual non-constant vectors \( \mathcal{V} \)

Recall that for \( r > 0, \delta > 0, \rho > 0 \) and a growth function \( g \),

\[
\mathcal{V}(r, g, \delta, \rho) = \left\{ x \in \mathcal{Y}(r) : \forall i \in [n], x_i^* \leq g(n/i) \text{ and } \exists Q_1, Q_2 \subset [n] \text{ such that } |Q_1|, |Q_2| \geq \delta n \text{ and } \max_{i \in Q_2} x_i \leq \min_{i \in Q_1} x_i - \rho \right\}.
\]

We will define our function \( g \) piecewisely on the intervals \([1, \frac{n}{n_{s+2}}), [\frac{n}{n_{s+2}}, \frac{n}{n_{s+1}}), [\frac{n}{n_{s+1}}, \frac{n}{n_s}), [\frac{n}{n_s}, \frac{n}{n_{s-1}}), \ldots, [\frac{n}{n_1}, \infty)\).

**Definition 3.5.** Let \( g \) be the growth function defined piecewisely on the intervals \([1, \frac{n}{n_{s+2}}), [\frac{n}{n_{s+2}}, \frac{n}{n_{s+1}}), [\frac{n}{n_{s+1}}, \frac{n}{n_s}), [\frac{n}{n_s}, \frac{n}{n_{s-1}}), \ldots, [\frac{n}{n_1}, \infty)\). First,

\[
g(t) = 2t^{3/2} \quad 1 \leq t \leq \frac{n}{n_{s+1}},
\]

\[
g(t) = 2t^3 \quad \frac{n}{n_{s+1}} \leq t \leq \frac{n}{n_s}.
\]

For \( j \in \{0, 1, 2, \ldots, s-2\} \), we define \( g_j : [1, \infty) \mapsto [1, \infty) \) by

\[
g(t) = \frac{t}{n/n_{s-j}}(\kappa pn)^j(pn)^4 \quad n/n_{s-j} \leq t < n/n_{s-j-1}
\]

and

\[
g(t) = \frac{t}{n/n_1}(\kappa pn)^{s-1}(pn)^4 \quad n/n_1 \leq t < \infty.
\]

In particular, \( g \) is a function depending on \( n, p, r, \kappa, \) and \( C_T \).

**Proposition 3.6.** There exists a universal constant \( K_3 > 0 \) such that the following holds: When \( n \) is sufficiently large, the growth function \( g \) from Definition 3.5 is a non-decreasing function and satisfies (14):

\[
\forall a \geq 2, \forall t \geq 1 \quad g(at) \geq g(t) + a \text{ and } \prod_{j=1}^{\infty} g(2^j)^{2^{s-j}} \leq K_3.
\]

Secondly, for \( x \in \mathcal{Y}_n(r) \setminus \mathcal{T} \),

\[
x_i^* \leq g\left(\frac{n}{i}\right)
\]

for \( \frac{n}{n_{s+1}} \leq i \leq n \).

Thirdly, we have \( \sum_{i \in [n]} g(i) = n^{1+o_n(1)} \).

**Proof.** To show that \( g \) is a non-decreasing function, it is sufficient to show that

\[
\lim_{i \to n/n_{s-j}} g(i) \leq g(n/n_{s-j}). \quad (21)
\]

for \( j \in \{-1, 0, 1, \ldots, s-1\} \), which is immediate when \( j = -1 \).
For \( j = 0 \),

\[
\lim_{i \to n/n_s} g(i) = 2(480en^3) \leq (pn)^4 \leq g\left(\frac{n}{n_s}\right)
\]

where we used \( n_s = \frac{1}{480e} \) in the first equality. For \( j \in [s - 1] \), we have \( \frac{n_{s-j+1}}{n_{s-j}} \leq pn \leq \kappa pn \) and therefore

\[
\lim_{i \to n/n_{s-j}} g(i) = \frac{n_{s-j+1}}{n_{s-j}}(\kappa pn)^{j-1}(pn)^4 \leq g\left(\frac{n}{n_{s-j}}\right).
\]

Hence, we conclude that \( g \) is a non-decreasing function.

Next, observe that \( g \) satisfies \( g(at) \geq g(t) \) with \( a \geq 1 \) in each interval which \( g \) defined piecewisely. And this property automatically extended to the whole domain \([1, \infty)\) due to (21). We conclude that for \( a \geq 2 \) and \( t \geq 1 \),

\[
g(at) \geq ag(t) \geq g(t) + a,
\]

where the second inequality relies on \( g(t) \geq 2 \) for \( t \geq 1 \). For the second condition from (14), it is not hard to see that \( g \) is growing like a polynomial. Specifically, one can easily argue that \( g(t) \leq 2t^{10} \). Thus, for \( j \geq 1 \),

\[
\log(g(2^j)^{2^{-j}}) \leq 11j^{2}2^{-j}
\]

and therefore, \( \sum_{j=1}^{\infty} g(2^j)^{2^{-j}} \leq K_3 \) for some sufficiently large \( K_3 > 0 \).

It remains to show the comparison \( x_i^* \leq g\left(\frac{\pi}{r}\right) \) for \( x \in Y(r) \setminus T \).

First of all, by Proposition 3.4

\[
x_i^* \leq C\frac{\sqrt{\kappa}}{\sqrt{T}}pn.
\]

For \( n_s \leq i \leq n_{s+1} \),

\[
g\left(\frac{n}{i}\right) \geq g\left(\frac{n}{n_{s+1}}\right) = g\left(\sqrt{pn}\right) \geq 2(pn)^{3/2} \geq C\frac{\sqrt{\kappa}}{\sqrt{T}}pn \geq x_i^*.
\]

The argument for \( n_{s-j} \leq i < n_{s-j+1} \) and \( j \in [s] \) is similar. By Proposition 3.3, \( x_{n_s-j}^* \leq C\frac{\sqrt{\kappa}}{\sqrt{T}}pn(\kappa pn)^j \).

Thus,

\[
g\left(\frac{n}{i}\right) \geq g\left(\frac{n}{n_{s-j+1}}\right) = (\kappa pn)^{j-1}(pn)^4 \geq x_{n_s-j}^* \geq x_i^*.
\]

We need to show the next Theorem

**Theorem 3.7.** For sufficiently large \( n \), we have the following:

Let \( r \in (0, \frac{1}{10}) \), \( \delta \in (0, \frac{\pi}{3}) \), \( \rho \in (0, \frac{1}{10}) \). Then,

\[
\mathbb{R}^n \setminus (\bigcup_{\lambda > 0} \lambda Y_n(r, g, \delta, \rho)) \subset (\bigcup_{\lambda > 0} \lambda \mathcal{R}) \cup T \cup \{\vec{0}\}
\]

where \( g \) is the growth function from Definition 3.5.
Proof. First of all, observe that if \( x \notin \cup_{\lambda > 0} \mathcal{Y}(\ell) \), \( x \in T \cup \{0\} \). Now we only need to consider \( x \in \cup_{\lambda > 0} \mathcal{Y}(\ell) \). In this case, it is sufficient to show

\[
\mathcal{Y}(\ell) \setminus \mathcal{V}(\ell, g, \delta, \rho) \subset R \cup T.
\]

Let \( x \in \mathcal{Y}(\ell) \setminus \mathcal{V}(\ell, g, \delta, \rho) \cup T \). It is sufficient to show \( x \in R \). Suppose there exists no \( Q_1, Q_2 \subset [n] \) with \( |Q_1|, |Q_2| \geq \delta n \) such that \( \max_{i \in Q_2} \min_{j \in Q_1} x_i - x_j \leq \rho \). Then, there exists a subset \( I \) of size \( n - 2\delta n \) such that for \( i, j \in I, |x_i - x_j| \leq \rho \). In particular, \( \sigma_x([rn]) \subseteq I \). Therefore, \( x \in \mathcal{A}(\rho) \).

As a consequence, we have \( \|x^*_{[0, n]}\| \geq \sqrt{(n - \lfloor rn \rfloor - 2\delta n)(1 - \rho)} \geq \frac{1}{2}\sqrt{n} \) where we used \( r, \rho < \frac{1}{10} \) and this provides us the bound that \( \|x^*_{[n, m]}\| \leq C_T \sqrt{\rho m} \sqrt{n} \) for \( n_{s+1} \leq k \leq \lfloor rn \rfloor \).

Let \( m_0 := \frac{n}{\log^2(pn)} \). Notice that we have \( m_0 \geq 2n_{s+1} \) when \( n \) is sufficiently large. Consider the value of \( x^*_{m_0} \):

Case 1: \( x^*_{m_0} \leq \log^2(pn) \)

\[
\|x^*_{[m_0, n]}\|_\infty \geq \frac{\sqrt{n}}{2 \log^2(pn)} \geq \frac{2C_{R_{g\bar{z}}} \sqrt{n}}{\sqrt{p}},
\]

Case 2: \( x^*_{n_{s+1}} \geq \log^2(pn) \)

\[
\|x^*_{[n_{s+1}, n]}\|_\infty \geq \sqrt{m_0 - n_{s+1} \log^2(pn)} \geq \frac{\sqrt{n} \log(pn)}{\sqrt{2}} \geq \frac{2C_{R_{g\bar{z}}} \sqrt{n}}{\sqrt{p}}.
\]

In the first case, \( x \in R^1_{m_0} \), and in the second case, \( x \in R^1_{n_{s+1}} \). Hence, \( x \in R \).

Now, suppose \( x^*_{i_0} > g\left(\frac{n}{i_0}\right) \) for some \( i_0 \in [n] \). If there are multiple indices satisfying this inequality, let \( i_0 \) be the smallest index. First of all, since \( x^*_{[n]} = 1, i_0 \leq \lfloor rn \rfloor \). Also, by Proposition 3.6, \( i_0 > n_{s+1} \).

For \( n_{s+1} < i_0 \leq rn \), we will show that if such \( x \) exists, then \( x \in R \). First of all, for \( k \leq \frac{i_0}{2} \),

\[
\|x^*_{[k, n]}\| \geq \|x^*_{[i_0, i_0]}\| \geq \sqrt{\frac{i_0}{2} g(n/i_0)} = \sqrt{\frac{i_0}{2} (2\frac{n}{i_0})^{3/2}} \geq \frac{2}{r} \sqrt{n},
\]

and notice that the last term is the lower bound of \( \|y_{[k, n]}^*\| \) for \( y \in R^1_k \). It remains to show that \( \|x^*_{[k, n]}\|_\infty \geq \frac{2C_{R_{g\bar{z}}} \sqrt{n}}{\sqrt{p}} \) for some \( n_{s+1} \leq k \leq \frac{n}{\log^2(pn)} \).

Suppose \( n_{s+1} \leq i_0 \leq \frac{n}{2 \log^2(pn)} \), let \( k = \frac{i_0}{2} \geq \frac{1}{2} n_{s+1} \geq n_{s+1} \). We have

\[
\|x^*_{[k, n]}\|_\infty \geq \sqrt{\frac{i_0}{2} g\left(\frac{n}{i_0}\right)} = 2 \sqrt{i_0} \geq \sqrt{\frac{n}{p}} \geq \frac{2C_{R_{g\bar{z}}} \sqrt{n}}{\sqrt{p}}.
\]

In the case \( 2 \frac{n}{\log^2(pn)} \leq i_0 \leq rn \), let \( k = \frac{n}{\log^2(pn)} \). We have

\[
\|x^*_{[k, n]}\|_\infty \geq \sqrt{\frac{i_0}{2} g\left(\frac{n}{i_0}\right)} \geq \sqrt{\frac{i_0}{2} \left(\frac{k}{i_0}\right)^{3/2}} \geq \frac{k^{3/2}}{\sqrt{2r}} \geq \frac{\sqrt{n}}{\sqrt{2r \log^3(pn)}} \geq \frac{2C_{R_{g\bar{z}}} \sqrt{n}}{\sqrt{p}}.
\]

\( \square \)
4 \quad T_1 \text{ vectors}

Let \( \Omega_0 \) be the event that \( |\mathcal{L}_A| = 0 \). In this section, we will prove the following theorem

**Theorem 4.1.** Assuming that \( \kappa > 1 \) is sufficiently large, there exists an event \( \Omega \), depending on \( p \) and \( n \), with \( \mathbb{P} \{ \Omega^c \} \leq 5n \exp(-\frac{6}{5}pn) \) such that condition on \( \Omega \cap \Omega_0 \), the following holds: For any \( x \in T_1 \),

\[
|\mathfrak{A}x| \geq n^{-1-o(1)} \|x\|.
\]

The proofs are not exactly the same when \( \tau > 10 \) or \( 1 \leq \tau \leq 10 \). By the definition of \( T \) and Proposition 3.3, Theorem 4.1 is a direct consequence of the following two theorems:

**Theorem 4.2.** For \( 1 \leq \tau \leq 10 \), there exists an event \( \Omega \), depending on \( p \) and \( n \), with \( \mathbb{P} \{ \Omega^c \} \leq 5n \exp(-\frac{6}{5}pn) \) such that conditioning on \( \Omega \cap \Omega_0 \), the following holds: For any \( x \in \mathbb{R}^n \) with \( x_{n_1}^* > \kappa pn x_{n_2}^* \) where \( \kappa > 1 \) is a sufficiently large constant and \( n_1, n_2 \) satisfies one of the following:

1. \( 1 \leq n_1 \leq \left\lceil \exp\left(\frac{pn}{\log^2(pn)}\right) \right\rceil \quad n_1 \leq n_2 \leq 3n_1 \)
2. \( \left\lceil \exp\left(\frac{pn}{\log^2(pn)}\right) \right\rceil \leq n_1 \leq \frac{1}{480e}\frac{pn}{\log^3(pn)}n_1, \frac{1}{480e}\frac{pn}{\log^3(pn)}n_1 \)

Then, we have

\[
|\mathfrak{A}x| \geq \frac{1}{4} x_{n_1}^*.
\]

where \( C > 0 \) is a universal constant.

**Theorem 4.3.** For \( 10 \leq \tau < +\infty \), there exists an event \( \Omega \), depending on \( p \) and \( n \), with \( \mathbb{P} \{ \Omega^c \} \leq 5n \exp(-\frac{6}{5}pn) \) such that conditioning on \( \Omega \cap \Omega_0 \), the following holds: For any \( x \in \mathbb{R}^n \) with \( x_{n_1}^* > 24pn x_{n_2+1}^* \) where \( n_1, n_2 \) satisfies one of the following:

1. \( n_1 = 1 \quad n_1 \leq n_2 \leq 2 \)
2. \( 2 \leq n_1 \leq \frac{1}{480e}\frac{pn}{\log^3(pn)}n_1 \)

Then, we have

\[
|\mathfrak{A}x| \geq \frac{1}{4} x_{n_1}^*.
\]

The event \( \Omega \) described in the above theorems are based on the statistics of supports of \( \{C_iA\}_{i \in [n]} \). Let us begin with the setup. Recall the definition of \( \mathcal{L}_A(k) \) from (4), let

\[
\mathcal{L}(k) := \mathcal{L}_A(k) = \{ j \in [n] : |\{ i \in [n] : a_{ij} = 1 \}| \leq k \}.
\]

Next, we will define a high probability event for \( \mathfrak{A} \) on \( \mathcal{L}(k) \) for different values of \( k \):

For \( 1 \leq k \leq \frac{\tau}{2}pn \), let

\[
u = \nu(k) := e2 \left(\frac{3pn}{k}\right)^k \exp(-pn)n
\]

and define

\[
L_k := \left\{ \begin{array}{ll}
\left\lceil \frac{\frac{6}{5}pn}{\log(\frac{2}{e})} \right\rceil & u < \frac{1}{e}, \\
\max\left\{ \frac{\frac{6}{5}pn}{\log(\frac{2}{e})}, eu \right\} & u \geq \frac{1}{e}.
\end{array} \right.
\]

Let \( \Omega_1 \) be the event that
For $1 \leq k \leq \frac{1}{2}pn$, $|\mathcal{L}(k)| < L_k$, and

$|\mathcal{L}(3pn)| = [n]$.

Also, let $\Omega_{\text{row}}$ be the event that the support size of every row of $\mathcal{A}$ does not exceed $3pn$.

**Lemma 4.4.** We have the following probability bound

$$P\{\Omega^c_1\} \leq 2n \exp(-\frac{6}{5}pn)$$

and

$$P\{\Omega^c_{\text{row}}\} \leq n \exp(-\frac{6}{5}pn).$$

**Proof.** First, by (9), for $j \in [n]$ we have $P\{j / \in \mathcal{L}(3pn)\} \leq \exp(-\frac{6}{5}pn)$. Next, by an union bound argument, $P\{\mathcal{L}(3pn) \neq [n]\} \leq n \exp(-\frac{6}{5}pn)$.

Fix $1 \leq k \leq \frac{1}{2}pn$, we want to choose $L_k > 0$ such that $P\{|\mathcal{L}(k)| \geq L_k\} \leq \exp(-\frac{6}{5}pn)$. Then, applying the union bound argument we obtain the statement of the lemma.

By (7), for $j \in [n],$

$$P\{j \in \mathcal{L}(k)\} \leq 2 \left(\frac{3pn}{k}\right)^k \exp(-pn) = \frac{u}{en}.$$ 

Next, since columns of $\mathcal{A}$ are independent, by (6), for $s > 2un$, we have

$$P\{|\mathcal{L}(k)| \geq s\} \leq \left(\frac{u}{s}\right)^s.$$ 

When $u < \frac{1}{e}$, $L_k = \lceil \frac{6pn}{\log(u)} \rceil$ and thus

$$P\{|\mathcal{L}(k)| \geq L_k\} \leq u^{L_k} = \exp(-\frac{6}{5}pn).$$

When $u \geq \frac{1}{e}$, $L_k = \lceil \max\{\frac{6pn}{5}, eu\} \rceil$ and thus

$$P\{|\mathcal{L}(k)| \geq L_k\} \leq \exp(-\frac{6}{5}pn).$$

By the union bound argument, we have

$$P\{\Omega^c_1\} \leq 2n \exp(-\frac{6}{5}pn) .$$

Due to $\mathcal{A}$ and $\mathcal{A}^\top$ have the same distribution, the proof of the probability estimate for $\mathcal{L}(3pn) = [n]$ is the proof of the estimate of probability for $\Omega^c_{\text{row}}$. 

$\square$
Let $k$ be a positive integer which we will specify later and let $C \subset [n]$ be the set of column indices

$$C := \{ j \in [n] : |\text{supp}(C_j(A))| \leq k \} = \mathcal{L}(k).$$

Let $R \subset [n]$ be the set of row indices

$$R := \{ i \in [n] : \exists j \in C \text{ s.t. } a_{ij} = 1 \}.$$

We could split the matrix $A$ into the block forms:

$$\begin{pmatrix}
\mathcal{H} & \mathcal{W} \\
0 & \mathcal{D}
\end{pmatrix}$$

where

1. submatrix $\mathcal{H}$ corresponds to rows $R$ and columns $C$,
2. submatrix $\mathcal{W}$ corresponds to rows $R$ and columns $[n]\setminus C$,
3. submatrix $\mathcal{D}$ corresponds to rows $[n]\setminus R$ and columns $[n]\setminus C$,

and the last submatrix is the one corresponding to rows $[n]\setminus R$ and columns $C$ which is a submatrix with all 0 entries.

We will describe some high probability events corresponding to these submatrices. And if we condition on these events, then $|Ax|$ is non-zero for every steep vector $x$.

We organize this section in the following way: The key tool Expansion Property will be described in the next subsection. Then, we need to split the remaining parts for $1 \leq \tau \leq 10$ and $10 \leq \tau < \infty$.

### 4.1 Expansion Property

For a $m_1 \times m_2$ Bernoulli $(p)$ random matrix $A$, and $\bar{b} = (b_1, \ldots, b_{m_2}) \in \{0, 1, \ldots, m_1\}^{m_2}$, let $\Omega_{\bar{b}}(A)$ be the event that $|\text{supp}(C_i(A))| = b_i$ for $i \in [m_2]$. Notice that, conditioning on $\Omega_{\bar{b}}(A)$, $\{C_i(A)\}_{i \in [m_2]}$ are jointly independent. For each $i \in [m_2]$, the support of $C_i(A)$ is chosen uniformly among subsets of $[m_1]$ with size $b_i$.

For $J_1 \subset J_2 \subset [m_2]$, we define

$$I_A(J_1, J_2) := \{ i \in [m_1] : \exists j_0 \text{ s.t. } a_{ij_0} = 1 \text{ and } a_{ij} = 0 \text{ for } j \in J_2 \setminus \{j_0\} \}.$$

**Lemma 4.5.** Let $A$ be an $m_1 \times m_2$ Bernoulli $(p)$ matrix and condition on the event $\Omega_{\bar{b}}(A)$ where $\max_{i \in m_2} b_i \leq b$. Let $J_1 \subset J_2 \subset [m_2]$ be two subsets. Suppose there exists $r > 0$ such that $\{ j \in J_1 : b_j \geq r \} \geq \frac{|J_1|}{2}$, and

$$\frac{|J_2|}{r} \geq \frac{2b_2^{2/m_1}}{m_1}.$$ 

Then,

$$\mathbb{P} \{ I(J_1, J_2) = 0 \mid \Omega_{\bar{b}}(A) \} \leq C_{1/2}^{\frac{|J_1|}{h_2}} \exp \left( - \log \left( \frac{r}{24|J_2|b_2^2/m_1} \right) \frac{r|J_1|}{8} \right).$$

The proof of this lemma relies on the following Proposition:

**Proposition 4.6.** Let $T_0, T_1, \ldots, T_k$ be independent random subsets of $[m_1]$ where
1. $T_0$ is uniformly chosen among all subsets of $m_1$ of size $s_0$.

2. $T_j$ is uniformly chosen among all subsets of $m_1$ of size $s_j$ where $1 \leq s_j \leq s$.

Consider their union $\bar{T}_i = (\cup_{s=0}^i T_s)$ and $X_i := |T_i \cap \bar{T}_{i-1}|$ for $l \in [k]$. Let $S = \sum_{j=0}^{k} s_j$. For $t > \max\{\frac{6S}{m_1}, 1\}$, we have

$$\mathbb{P}\left\{ \sum_{i=1}^{k} X_i \geq tk \right\} \leq C_{\text{hg}}^k \exp\left( -\log\left( \frac{t}{6S/m_1} \right) tk \right).$$

**Proof.** For simplicity, we assume $|J_1| = n_1$ and $|J_2| = n_2$. Without lose of generality, we assume that $[k] \subset J_1$, $k \geq \frac{n_1}{2}$, and $b_i > r$ for $i \in [k]$.

For $j \in [k]$, let $T_j$ be the support of $C_j(A)$. And $T_0$ be the union of supports of $C_j(A)$ for $j \in J_2 \setminus [k]$. Condition on $\Omega_0(A)$, we have $|T_0| \leq (n_2-k) \cdot b$. Next, we condition on $|T_0|$, then we could apply Proposition 4.6 with $t = \frac{r}{4}$ to get:

$$\mathbb{P}\left\{ \sum_{i=1}^{k} X_i \geq \frac{rn_1}{8} \right\} \leq C_{\text{hg}}^{n_1/2} \exp\left( -\log\left( \frac{r}{24n_2b^2/m_1} \right) \frac{rn_1}{8} \right).$$

Now, let $I_i = \{i : \exists j \in [l] \text{ s.t. } a_{ij} = 1 \text{ and } a_{ij} = 0 \text{ for } j \in T_0\}$. Observe that $|I_1| = |T_1| - X_1$ and recursively we have $|I_j| \geq |I_{j-1}| + |T_j| - 2X_j$. Combining these inequality together we get

$$|I_A(J_1, J_2)| \geq |I_k| \geq \sum_{t \in [k]} |T_t| - 2 \sum_{t \in [k]} X_j \geq \frac{kr}{2} - 2 \sum_{t \in [k]} X_j \geq \frac{n_1r}{4} > 0.$$

$\square$

**Proof of Proposition 4.6** First, $\{X_i\}_{i \in [k]}$ is not jointly independent, we will first resolve this issue.

For $l \in [k]$, let $\sigma_l$ be a permutation on $[n] \setminus \mathcal{L}$ determined by $\bar{T}_{l-1}$ satisfying $\sigma_l([|\bar{T}_{l-1}|]) = \bar{T}_{l-1}$. Observe that $(T_0, T_1, \ldots, T_k)$ and $(T_0, \sigma_1(T_1), \ldots, \sigma_k(T_k))$ have the same distribution. (which can be verified inductively.)

In particular, now we can view $X_l = |[|\bar{T}_{l-1}|] \cap T_l|$. Let $Y_l = |\sum_{j=0}^{k} |T_j|| \cap T_l|$. Then, we have $X_l \leq Y_l$ and $\{Y_j\}_{j=1}^{k}$ are jointly independent. We will bound $\sum_{j=1}^{k} Y_j$ instead.

Let $S = \sum_{j=0}^{k} s_j$. By $(\text{II})$,

$$\mathbb{P}\{Y_j = u\} \leq C_{\text{hg}} \left( \frac{3Ss}{um_1} \right)^u.$$

Let $u > \max\{\frac{3Ss}{m_1}, 1\}$, we have $\lambda(u) := \log\left( \frac{um_1}{3Ss} \right) > 0$. For $t \geq 1$, we have

$$\mathbb{P}\{Y_j = t\} \exp(\lambda(u)t) \leq C_{\text{hg}} \left( \frac{u}{t} \right)^t.$$

Thus,

$$\mathbb{E}\exp(\lambda(u)Y_j) \leq 1 + \sum_{t=1}^{\infty} C_{\text{hg}} \left( \frac{u}{t} \right)^t.$$

To estimate the sum, notice that the function $t \rightarrow \left( \frac{u}{t} \right)^t$ reaches its maximum when $t = \frac{u}{e}$ with value $e^{\frac{u}{e}}$. And for $t \geq 2u$, it is bounded by $\left( \frac{u}{2} \right)^t$. Therefore,
\[ \mathbb{E} \exp(\lambda(u)Y_j) \leq 1 + 2uC_{\text{hg}} \exp(u/e) + 2C_{\text{hg}} \left( \frac{1}{2} \right)^{2u} \leq C_{\text{hg}} \exp(\frac{2}{e}u). \]

Therefore, we have

\[
P\left\{ \sum_{j=1}^{k} X_j \geq tk \right\} \leq P\left\{ \sum_{j=1}^{k} Y_j \geq tk \right\} \leq \mathbb{E} \exp(\lambda(u) \sum_{j=1}^{k} Y_j) \exp(-\lambda(u)tk) \leq C_{\text{hg}}^k \exp(\frac{2}{e}uk - \lambda(u)tk).\]

The value \( \frac{2}{e}uk - \lambda(u)tk \) is minimized when \( \frac{2}{e}u = t \). Thus, for \( t \geq \max\{\frac{6Ss}{m_1}, 1\} \), we have

\[
P\left\{ \sum_{j=1}^{k} X_j \geq tk \right\} \leq C_{\text{hg}}^k \exp(-\log(\exp(\frac{2}{3}\frac{t}{Ss/m_1})tk) = C_{\text{hg}}^k \exp(-\log(\frac{t}{6Ss/m_1})tk).\]

\[ \square \]

### 4.2 Typical \( \mathcal{A} \) for \( 1 \leq \tau \leq 10 \)

In this regime, \( \mathcal{A} = (a_{ij})_{i,j \in [n]} \) is a \( n \times n \) Bernoulli(\( p \)) matrix and \( \tau = \frac{pn}{\log(n)} \) satisfies

\[ 1 \leq \tau \leq 10. \]

Here we will evaluate the values of \( L_r \) for a few specific \( r \) which will be used regime.

Let \( t \geq 1 \) be a fixed large constant which will be determined later.

**Proposition 4.7.** For \( 1 \leq \tau \leq 10 \), we have

\[ L_t \leq \frac{1}{100} \log^2(n), \quad L_{\frac{100pn}{\log^3(pn)}} \leq \frac{1}{100} \exp\left( \frac{pn}{\log^2(pn)} \right), \quad \text{and} \quad L_{\frac{1}{2}pn} \leq \frac{1}{100} n^{0.9}. \quad (24) \]

**Proof.** The estimate is straightforward from the definition of \( u \) from (22) and the definition of \( L_r \) from (23).

For \( r = t \) which is a constant,

\[ u(t) \leq 2e(3pn)^t \leq (\log(n))^{3/2}, \]

which holds when \( n \) is sufficiently large.

For \( r = \frac{100pn}{\log^3(pn)} \),

\[ u = 2e\left( \frac{3\log^3(pn)}{100} \right) \frac{100pn}{\log^3(pn)} \exp(-pn)n \]
\[ \leq 2e \exp\left( -(1 - \frac{3\log(\log(pn))}{\log^3(pn)})pn + \log(n) \right) \]
\[ \leq \exp\left( \frac{\log(n)}{\log^{5/2}(\log(n))} \right), \]

where we used the fact that \( pn \geq \log(n) \).
For $r = \frac{1}{2}pn$, similarly we have

\[
\begin{align*}
u &= 2e(6)^{\frac{1}{2}pn}\exp(-pn)n \\
&\leq 2e \exp \left( \frac{\log(6)}{2}pn - pn + log(n) \right) \\
&\leq 2e \exp \left( \left( \frac{9}{10} - \frac{1}{1000} \right)\log(n) \right),
\end{align*}
\]

where we used $\frac{\log(6)}{2} < \frac{9}{10} - \frac{1}{1000}$.

Notice that by [23], $u \mapsto L_r$ is a monotone increasing function. By applying $L_r = eu$ we obtained the desired inequalities.

We will described typical events of $\calH$, $\calW$, and $\calD$.

Let $\Omega_\calC$ be the event that for $j_1, j_2 \in \calC$, $\text{supp}(C_{j_1}(A)) \cap \text{supp}(C_{j_2}(A)) = \emptyset$.

Let $\Omega_{\calW}$ be the event that $|\text{supp}(C_j(\calW))| \leq 30$ for all $j \in [n] \setminus \calC$.

Let $\Omega_\calD$ be the event that

- For $1 \leq n_1 \leq \lceil \exp(\frac{pn}{\log^2(pn)}) \rceil$ and any subsets $J_1 \subset J_2 \subset [n] \setminus \calC$ with $|J_1| = n_1$ and $|J_2| \leq 1000n_1$, $|I_\calD(J_1, J_2)| > 0$.

- For $\frac{1}{10} \lceil \exp(\frac{pn}{\log^2(pn)}) \rceil \leq n_1 \leq \frac{1}{480e} \lceil \exp(\frac{pn}{\log^2(pn)}) \rceil$, and any subsets $J_1 \subset J_2 \subset [n] \setminus \calC$ with $|J_1| = n_1$ and $|J_2| \leq \max\{\frac{1}{480e}, 3\frac{pn}{\log^2(pn)}n_1\}$, $|I_\calD(J_1, J_2)| > 0$.

In the end, let $\Omega_2 = \Omega_1 \cap (\Omega_\calC \cap \Omega_{\calW} \cap \Omega_\calD) \cap \Omega_{\text{row}}$.

We will show that the probability of $\Omega_2$ is insignificant comparing to that of $\Omega_0$. And for a sample $A$ of $\frakA$ from $\Omega_0 \cap \Omega_2$, $|Ax|$ is non-zero for every steep vector $x \in \calT$.

Theorem 4.2 will be break into the following lemma and theorem:

**Lemma 4.8.** The following probability estimate holds:

\[
P\{\Omega_2^c\} \leq 6n \exp(-\frac{6}{5}pn).
\]

**Theorem 4.9.** Fix a sample of $A$ from $\Omega_0 \cap \Omega_2 \cap \Omega_{\text{row}}$. Consider $x \in \mathbb{R}^n$ such that $x^*_{n_1} \geq \kappa pn x^*_{n_2+1}$ where $\kappa > 0$ is a universal constant and $n_1, n_2$ satisfies one of the following:

1. $1 \leq n_1 \leq \lceil \exp(\frac{pn}{\log^2(pn)}) \rceil$ \hspace{1cm} $n_1 \leq n_2 \leq 3n_1$

2. $\lceil \exp(\frac{pn}{\log^2(pn)}) \rceil \leq n_1 \leq \frac{1}{480e} \lceil \exp(\frac{pn}{\log^2(pn)}) \rceil$ \hspace{1cm} $n_1 \leq n_2 \leq \min\{\frac{pn}{\log^3(pn)}n_1, \frac{1}{480e}\}$.

Then, we have

\[
|Ax| \geq 1.9 p^*_{n_1}
\]

where $1.9 > 0$ is a universal constant.
4.2.1 Probability Estimate for $\Omega^5_\mathfrak{A}$.

To prove Lemma 4.8 we need to get estimates of probability events related to $\Omega_\mathcal{C}$, $\Omega_\mathcal{R}$, and $\Omega_\mathfrak{D}$.

**Proposition 4.10.** In the regime $\tau \leq 10$, the following bound for the probability holds:

$$\mathbb{P}\{\Omega^c_\mathcal{C} \mid \Omega_1\} \leq n^{-0.9} \leq n \exp(-1.2pn).$$

**Proof.** We condition on a subevent $\Omega^5_\mathfrak{A} \subset \Omega_1$ By, the probability that the intersection of the supports of any two columns corresponding to $\mathcal{C}$ is non-empty is bounded by $C_1 \frac{3(\ell)^2}{n}$. Since $\mathfrak{A}$ is a fixed constant, by (23),

$$|\mathcal{C}| \leq \sum_{k=0}^\ell L_k \leq \mathfrak{A} 2e^2 \left(\frac{3pn}{\ell}\right) \leq \log(n)^2.$$

The Proposition is proved by applying the union bound argument. \qed

**Lemma 4.11.** The following bound for probability holds:

$$\mathbb{P}\{\Omega^c_\mathcal{R} \mid \Omega_1\} \leq \exp(-2pn).$$

**Proof.** We condition on a subevent $\Omega^5_\mathfrak{A}$ of $\Omega_1$. Furthermore, we condition on $\mathcal{C} = \mathcal{C}_0, \mathcal{R} = \mathcal{R}_0$.

Notice that, we still have $\{C_j(\mathfrak{A})\}_{j \in [n]_\mathcal{C}}$ are jointly independent, and $\text{supp}(C_j(\mathfrak{A}))$ is chosen uniformly among subsets of $[n]$ with size $b_j$ for $j \in [n] \setminus \mathcal{C}$.

On the other hand, $\mathcal{R}$ is a fixed set. According to the hypergeometric distribution, the expected size of $\mathcal{R} \cap \text{supp}(C_j(\mathfrak{A}))$ is $\frac{|\mathcal{R}| \|\text{supp}(C_j(\mathfrak{A}))\|}{n}$.

Since it is a subevent of $\Omega_1$, we have $|\mathcal{R}| \leq \mathfrak{A} |\mathcal{C}| \leq \mathfrak{A} \log^2(n)$ and $|\text{supp}(C_j(\mathfrak{A}))| \leq 3pn$.

Then, by the Hypergeometric distribution tail, for $j \in [n] \setminus \mathcal{C}$,

$$\mathbb{P}\{|\mathcal{R} \cap \text{supp}(C_j(\mathfrak{A}))| \geq l \mid \Omega^5_\mathfrak{A}, \mathcal{R} = \mathcal{R}_0, \mathcal{C} = \mathcal{C}_0\} \leq n^{-0.9l}.$$ (25)

Using $pn \leq 10 \log(n)$ and choosing $l = 30$, the union bound argument gives

$$\mathbb{P}\{\exists j \in [n] \setminus \mathcal{C} \mid |\mathcal{R} \cap \text{supp}(C_j(\mathfrak{A}))| \geq 6 \mid \Omega^5_\mathfrak{A}, \mathcal{R} = \mathcal{R}_0, \mathcal{C} = \mathcal{C}_0\} \leq \exp(-2pn).$$

Furthermore, this conditional probability holds for all realizations $\Omega^5_\mathfrak{A}, \mathcal{C}$, and $\mathcal{R}$ whenever it is a subevent of $\Omega_1$. Thus, the proof is finished. \qed

**Lemma 4.12.**

$$\mathbb{P}\{\Omega^c_\mathfrak{D} \mid \Omega_1 \cap \Omega_2\} \leq \exp(-2pn).$$

**Proof.** We condition on $\mathcal{C} = \mathcal{C}_0, \mathcal{R} = \mathcal{R}_0$, and $\Omega^5_\mathfrak{D}$ which is a sub-event of $\Omega_1 \cap \Omega_2$.

For positive integers $n_1 \leq n_2 \leq n \setminus |\mathcal{C}|$,

$$|\{J_1 \subset J_2 \subset [n] \setminus \mathcal{C} : |J_1| = n_1 \text{ and } |J_2| = n_2\}| \leq \binom{n_2}{n_1} \binom{n}{n_2} \leq 2^{n_2} \left(\frac{en}{n_2}\right)^{n_2} \leq \exp\left(\log\left(\frac{2en}{n_2}\right)^{n_2}\right).$$

First, $\mathfrak{D}$ is a $n - |\mathcal{R}_0|$ by $n - |\mathcal{C}_0|$ and it is conditioned on $\Omega^5_\mathfrak{D}$.

Let $J_1 \subset J_2 \subset [n] \setminus \mathcal{C}$ with $|J_1| = n_1, |J_2| = n_2$.

Next, we want to apply Lemma 4.13 to give a probability estimate $I_\mathfrak{D}(J_1, J_2) = \emptyset$. 25
As a subevent of $\Omega_1$, we have
\[
n - |R_0| \geq \frac{9}{10} n, \quad \frac{r}{2} \leq \min_{i \in [n] \setminus C} b_i \leq \max_{i \in [n] \setminus C} b_i \leq 3pn.
\]

We apply Lemma 4.5 with a compatible choice of $r$, we have
\[
\begin{align*}
\mathbb{P}\left\{ I_D(J_1, J_2) = 0 \mid C = C_0, R = R_0, \Omega_b(D) \right\} \\
\leq C_{\text{hg}}^{n_1/2} \exp \left( - \log \left( \frac{r}{24n_2(3pn)^2} \right) \frac{rn_1}{16} \right) \\
= C_{\text{hg}}^{n_1/2} \exp \left( - \log \left( \frac{2en}{n_2} \frac{r}{480e(pn)^2} \right) \frac{rn_1}{16} \right).
\end{align*}
\]

Next, we will do the estimates depending on the particular values of $n_1$ and $n_2$.

Suppose $n_2 \leq \frac{n}{(pn)^{10}}$. Together with the assumption that $1 \leq r \leq pn$, we have
\[
1 \leq \frac{480e(pn)^2}{r} \leq \left( \frac{2en}{n_2} \right)^{0.1}
\]
and thus
\[
\log \left( \frac{2en}{n_2} \frac{r}{480e(pn)^2} \right) = \log \left( \frac{2en}{n_2} \right) - \log \left( \frac{480e(pn)^2}{r} \right) \leq \frac{9}{10} \log \left( \frac{2en}{n_2} \right).
\]

In this case, we obtain the probability bound
\[
\begin{align*}
\mathbb{P}\left\{ I_D(J_1, J_2) = 0 \mid C = C_0, R = R_0, \Omega_b(D) \right\} \\
\leq \exp \left( - \log \left( \frac{2en}{n_2} \right) \frac{rn_1}{32} \right).
\end{align*}
\]

Case 1: $1 \leq n_1 \leq \left\lceil \exp\left( \frac{pn}{\log^2(pn)} \right) \right\rceil$ and $n_2 \leq 1000n_1$.
Since we are conditioning on a subevent of $\Omega_{2W}$,
\[
\min_{i \in [n] \setminus C} b_i \geq \frac{r}{2} - 30 \geq \frac{r}{2},
\]
when $r \geq 60$. Thus, we could set $r = \frac{r}{2}$.

When $\frac{r}{2}$ is sufficiently large, $\frac{\frac{r}{64}}{2} > 2n_2$. Applying a union bound argument we get
\[
\mathbb{P}\left\{ \exists J_1 \subset J_2 \subset [n] \setminus C : |J_1| = n_1, |J_2| = n_2 \text{ and } I(J_1, J_2) = 0 \mid C = C_0, R = R_0, \Omega_b(D) \right\}
\leq \exp \left( - \log \left( \frac{2en}{n_2} \right) \frac{\frac{rn_1}{128}}{2} \right)
\leq \exp \left( - \frac{1}{2} \log(n) \frac{\frac{rn_1}{128}}{2} \right)
\leq \frac{1}{n^2} \exp(-2pn).
\]

Case 2: $\frac{1}{10} \left[ \exp\left( \frac{pn}{\log^2(pn)} \right) \right] \leq n_1 \leq \left\lceil \frac{n}{(pn)^{69}} \right\rceil$ and $n_2 \leq 3\frac{pn}{\log^{69}(pn)}n_1$. 

26
By (24), \( n_1 > 10L \frac{100\pi n}{\log^3(\pi n)} \), we can apply \( r = \lceil \frac{100\pi n}{\log^3(\pi n)} \rceil \). In particular, we have \( \frac{m_n}{10} \geq 2n_2 \).

\[
\mathbb{P} \{ \exists J_1 \subset J_2 \subset [n] \setminus \mathcal{C} : |J_1| = n_1, |J_2| = n_2 \text{ and } I(J_1, J_2) = 0 \mid \mathcal{C} = C_0, \mathcal{R} = R_0, \Omega_0(\mathcal{D}) \} \\
\leq \exp \left( -\log \left( \frac{2en}{n_2} \right) n_2 \right) \\
\leq \exp (-n_2) \\
\leq \frac{1}{n_2} \exp(-2\pi n),
\]

**Case 3:** \( \lceil \frac{n}{(\pi n)^{1/2}} \rceil \leq n_1 \leq \lceil \frac{1}{480e\pi} \rceil \) and \( n_2 \leq \max\{ \lceil \frac{1}{480e\pi} \rceil, 3 \frac{\pi n}{\log^3(\pi n)} n_1 \} \). By (24), \( n_1 > 100L \frac{\pi n}{\pi n} \), we can set \( r = \frac{1}{2} \pi n \). Now,

\[
\frac{2en}{n_2} \cdot \frac{\pi n}{480e(\pi n)^2} \geq e
\]

where \( n_2 \leq \frac{1}{480e\pi} \) is used. With these bounds, the individual probability estimate becomes

\[
\mathbb{P} \{ I_\mathcal{D}(J_1, J_2) = 0 \mid \mathcal{C} = C_0, \mathcal{R} = R_0, \Omega_0(\mathcal{D}) \} \\
\leq \exp \left( -\frac{rn_1}{16} \right) \\
\leq \exp (-\log^2(\pi n)n_2)
\]

Notice that in this case, \( \exp(\log(\frac{2\pi n}{n_2})n_2) \leq \exp(100 \log(\pi n)n_2) \). Hence,

\[
\mathbb{P} \{ \exists J_1 \subset J_2 \subset [n] \setminus \mathcal{C} : |J_1| = n_1, |J_2| = n_2 \text{ and } I(J_1, J_2) = 0 \mid \mathcal{C} = C_0, \mathcal{R} = R_0, \Omega_0(\mathcal{D}) \} \\
\leq C_{\text{hug}}^{n_1/2} \exp (-\log^2(\pi n)n_2 + 100 \log(\pi n)n_2) \\
\leq \exp (-n_2) \\
\leq \frac{1}{n_2} \exp(-2\pi n).
\]

Now, we are ready to prove Lemma 4.8

**Proof.** By definition,

\[
\Omega_c^c = \Omega_1^c \cup \Omega_2^c \cup \Omega_{2\pi} \cup \Omega_\delta \cup \Omega_{\text{row}}^c \\
= \Omega_1^c \cup (\Omega_2^c \cap \Omega_1) \cup (\Omega_{2\pi} \cap \Omega_1) \cup (\Omega_\delta \cap (\Omega_1 \cap \Omega_{2\pi}) \cup \Omega_{\text{row}}^c.
\]

By Proposition 4.10, Lemma 4.4, 4.11 and 4.12, we have

\[
\mathbb{P} \{ \Omega_1^c \} \leq 2n \exp(-\frac{6}{5} \pi n), \\
\mathbb{P} \{ \Omega_2^c \mid \Omega_1 \} \leq n \exp(-\frac{6}{5} \pi n), \\
\mathbb{P} \{ \Omega_{2\pi}^c \mid \Omega_1 \} \leq \exp(-2\pi n), \\
\mathbb{P} \{ \Omega_\delta^c \mid \Omega_1 \cap \Omega_{2\pi} \} \leq \exp(-2\pi n), \\
\mathbb{P} \{ \Omega_{\text{row}}^c \} \leq n \exp(-\frac{6}{5} \pi n).
\]


Combining these together we get
\[ \mathbb{P} \{ \Omega_2^c \} \leq 6n \exp(-\frac{6}{5}pn). \]

\[ \square \]

### 4.3 Estimate

Now we consider \( x \in \mathcal{T} \). We split in several cases. Let \( P_C \) be the orthogonal projection to the span of \( \{e_j\}_{j \in [n]} \setminus c \).

**Lemma 4.13.** Fix a sample of \( A \) from \( \Omega_0 \cap \Omega_2 \cap \Omega_{\text{row}} \). Let \( x \in \mathbb{R}^n \) be a vector such that there exists \( J \subset \mathcal{C} \) satisfying \( \min_{j \in J} |x_j| > \kappa pn x^*_0[J] \).

Then, either \( |Ax| > \frac{1}{4} \min_{j \in J} |x_j| \) or the vector \( y = P_C x \) satisfies
\[
y^*_\min(\frac{|y|}{\min_{j \in J} |x_j|}) \geq y^*_0[J], \quad y^*_\min(\frac{|y|}{\min_{j \in J} |x_j|}) \geq \frac{13}{4} \min_{j \in J} |x_j|
\]
where \( \kappa > 0 \) is a universal constant.

**Proof.** Let \( J_2 := \sigma_x(6|J|) \). Let \( \mathcal{R}(x) = \{ i \in \mathcal{R} : \exists j \in J \text{ s.t. } a_{ij} = 1 \} \).

For \( i \in \mathcal{R}(x) \), let \( j_0 \in J \) be an index such that \( a_{ij_0} = 1 \).

Since \( A \) is a sample in event \( \Omega_2 \subset \Omega_9 \), the existence of \( j_0 \) is unique. Thus, \( a_{ij} = 0 \) for \( j \in \mathcal{C}\setminus\{j_0\} \).

We have
\[
|Ax| = | \sum_{j \in [n]} a_{ij} x_j | \\
\geq |x_0| - | \sum_{j \in J_2 \setminus \mathcal{C}} a_{ij} x_j | - | \sum_{j \notin J_2 \setminus \mathcal{C}} a_{ij} x_j | \\
\geq (1 - \frac{3}{\kappa}) |x_{j_0}| - | \sum_{j \in J_2 \setminus \mathcal{C}} a_{ij} x_j | \\
\geq \frac{1}{2} \min_{j \in J} |x_j| - | \sum_{j \in J_2 \setminus \mathcal{C}} a_{ij} x_j |
\]
where the second inequality relies on \( |x_{j_0}| \geq \kappa pn x^*_0[J] \).

If \( | \sum_{j \in J_2 \setminus \mathcal{C}} a_{ij} x_j | \leq \frac{1}{4} |x^*_0| \), then
\[
|Ax| \geq \frac{1}{4} |x^*_0|.
\]

Now consider \( | \sum_{j \in J_2 \setminus \mathcal{C}} a_{ij} x_j | \geq \frac{1}{4} |x^*_0| \) for all \( i \in \mathcal{R} \).

For \( j \in J_2 \setminus \mathcal{C} \), since \( A \) is a sample from \( \Omega_1 \subset \Omega_9 \), \( |\mathcal{R}(x) \cap \text{supp}(C_J A)| \leq 30 \). Thus,
\[
|\{(i, j) \in \mathcal{R} \times (J_2 \setminus \mathcal{C}) : a_{ij} = 1\}| \leq 30|J_2|.
\]

There exists a subset \( \mathcal{R}(x)' \subset \mathcal{R}(x) \) with \( |\mathcal{R}(x)'| > \frac{1}{2} |\mathcal{R}| \) such that, for \( i \in \mathcal{R}(x) \), \( | \{ j \in J_2 \setminus \mathcal{C} : a_{ij} = 1 \} | \leq \frac{60|J_2|}{|\mathcal{R}(x)'|} \).

Due to \( A \) is a sample of \( \Omega_0 \), each column contains at least one non-zero entry. Thus, \( |\mathcal{R}(x)| \geq |J| \). We have \( | \{ j \in J_2 \setminus \mathcal{C} : a_{ij} = 1 \} | \leq 360 \).
Therefore, for each \( i \in \mathcal{R}(x)' \), we pick \( j_i \in J_2 \setminus C \) such that
\[
|x_{j_i}| \geq \frac{1}{360} \frac{1}{4} \min_{j \in J} |x_j| := c_{1.13} \min_{j \in J} |x_j|.
\]
Let
\[
J_1 := \{ j_i : i \in \mathcal{R}(x)' \}.
\]
Notice that it is possible that for \( i_1 \neq i_2, j_{i_1} = j_{i_2} \). As a subevent of \( \Omega_{2n} \), for \( j \in [n] \setminus C, \{ i \in \mathcal{R}(x)' : j_i = j \} \leq 30 \). Therefore, \( |J_1| \geq \min \{ \frac{|J|}{60}, 1 \} \).

Now, let \( y = Px \). We have \( |y_j| \geq \frac{1}{1.13} |x_{j_i}^*| \) for \( j \in J_1 \). And for \( j \notin J_2 \), we have \( |y_j| < \kappa pn |x_{j_i}^*| \). Now we conclude \( y_{\min \{ J_1, 1 \}}^* \geq 6pn y_{n+1}^* \).

**Proof of Theorem 4.9.** Now suppose \( 1 \leq n_1 \leq \lceil \exp(\frac{pn}{\log^* (pn)}) \rceil \) and \( n_1 \leq n_2 \leq 3n_1 \).

Suppose \( |\sigma_x([n_1]) \cap C| \geq \frac{n_1}{2} \). Let \( J = \sigma_x([n_1]) \cap C \). Then, \( n_2 \leq 6 |J| \). We apply Lemma 4.13 and one of the following is true: Either \( |Ax| > \frac{1}{2} |x_{n_1}^*| \), or \( y = P_c x \) satisfies
\[
y_{\min \{ J_1, 1 \}}^* > 6pn y_{n+1}^*, \quad \text{or} \quad y_{\min \{ J_1, 1 \}}^* > c_{1.13} y_{n_1}^*.
\]

There is nothing left to prove if \( |Ax| > \frac{1}{2} |x_{n_1}^*| \).

Let \( J_1 = \sigma_y([n_1]) \cap C \) and \( J_2 = \sigma_y([n_2]) \).

As a subevent of \( \Omega_{\Sigma} \), there exists \( i \in [n] \setminus \mathcal{R} \) such that \( i \in I(J_1, J_2) \). Let \( j_0 \in J_1 \) be the unique index such that \( a_{ij_0} = 1 \).

Then,
\[
|(Ax)_i| = |(Ay)_i| \geq |y_{\min \{ J_1, 1 \}}| - \sum_{j \notin J_2} a_{ij} y_j \geq y_{\min \{ J_1, 1 \}}^* - \frac{3pn}{6pn} y_{n_1}^* \geq \frac{1}{2} c_{1.13} y_{n_1}^*.
\]

where in the second inequality we used \( |\text{supp}(R_c(A))| \leq 3pn \) due to \( A \) is a sample from \( \Omega_{\text{row}} \).

In the case \( |\sigma_x([n_1]) \cap C| \leq \frac{n_1}{2} \), let \( y = P_c x \). Then, we have
\[
y_{\min \{ J_1, 1 \}}^* \geq \kappa pn y_{n+1}^*, \quad \text{or} \quad y_{\min \{ J_1, 1 \}}^* \geq x_{n_1}^*.
\]

Applying the same argument we obtain
\[
|Ax| \geq \frac{1}{2} x_{n_1}^*.
\]

Suppose \( \lceil \exp(\frac{pn}{\log^* (pn)}) \rceil \leq n_1 \leq \frac{1}{480e p} \) and \( n_1 \leq n_2 \leq \min \{ \frac{pn}{\log^* (pn)}, \frac{1}{480e p} \} \). Let \( y = P_c x \). Due to \( |C| \leq \frac{1}{2} \lceil \exp(\frac{pn}{\log^* (pn)}) \rceil \), we have \( y_{n_1/2}^* > \kappa pn y_{n+1}^* \) and \( y_{n_1/2}^* \geq x_{n_1}^* \).

Let \( J_1 = \sigma_y([n_1/2]) \) and \( J_2 = \sigma_y([n_2]) \). There exists \( i \in I(J_1, J_2) \) due to \( \Omega_{\Sigma} \). Let \( j_0 \in J_1 \) be the unique index such that \( a_{ij_0} = 1 \). By a similar comparison, we get
\[
|(Ax)_i| = |(Ay)_i| \geq \frac{1}{2} x_{n_1}^*.
\]

\[\square\]
4.4 \quad 10 \leq \tau < +\infty

The structure of the proof in this case is the similar to that of the case 1 \leq \tau \leq 10, but the technical details are different.

We will still use \mathcal{C}, \mathcal{R} and the corresponding submatrices \mathcal{H}, \mathcal{W}, and \mathcal{D}.

However, in this case, we set

\[ \ell = \frac{1}{100}pn. \]

**Proposition 4.14.** For \( \tau \geq 10 \), \( L_\ell = 2 \). As a consequence, condition on \( \Omega_1 \), the set \( \mathcal{C} = \mathcal{L}(\ell) \) satisfies

\[ |\mathcal{C}| \leq 1. \]

**Proof.** Recall the definition of \( u(r) \) and \( L(r) \) from (22) and (23). When \( r = \ell \),

\[ u = e \cdot 2 \exp\left( \frac{\log(300)}{100}pn - pn + \log(n) \right) \leq \exp\left( -\frac{4}{5}pn \right) \]

where we used \( \frac{\log(300)}{100} \leq 0.06 \) and \( \log(n) \leq \frac{1}{10}pn \).

Thus,

\[ L_\ell \leq \left\lfloor \frac{\frac{6}{5}pn}{\frac{3}{5}pn} \right\rfloor \leq \left\lfloor \frac{3}{2} \right\rfloor = 2. \]

\[ \square \]

Let \( \Omega_{2W} \) be the event that \( \forall j \in [n]\setminus\mathcal{C} \),

\[ |\text{supp}(\mathcal{C}_j(\mathcal{W}))| \leq \frac{1}{2} \ell. \]

Let \( \Omega_D \) be the event that for \( 1 \leq n_1 \leq \left\lfloor \frac{1}{480ep} \right\rfloor \), and any subsets \( J_1 \subset J_2 \subset [n]\setminus\mathcal{C} \) with \( |J_1| = n_1 \) and \( |J_2| \leq \max\{\left\lfloor \frac{1}{480ep} \right\rfloor, \frac{pn}{3\log^3(pn)}n_1\} \), then \( |I_D(J_1, J_2)| > 0 \).

Then, we define the event \( \Omega_2 = \Omega_1 \cap (\Omega_{2W} \cap \Omega_D) \cap \Omega_{row} \).

Similarly, the Theorem 4.3 will be break into the following lemma and theorem:

**Lemma 4.15.** The following probability estimate holds:

\[ \mathbb{P}\{\Omega_2\} \leq 6n \exp\left( -\frac{6}{5}pn \right). \]

**Theorem 4.16.** Fix a sample of \( A \) from \( \Omega_0 \cap \Omega_2 \cap \Omega_{row} \). Consider \( x \in \mathbb{R}^n \) such that \( x_{n_1}^* \geq 24pnx_{n_2+1}^* \) where we \( n_1, n_2 \) satisfies

| \begin{align*}
 1 & \quad n_1 = 1 & \quad n_1 \leq n_2 \leq 2,
 2 & \quad 2 \leq n_1 \leq \frac{1}{480ep} & \quad n_1 \leq n_2 \leq \max\{\left\lfloor \frac{1}{480ep} \right\rfloor, \frac{pn}{3\log^3(pn)}n_1\}
\end{align*} |

Then, we have

\[ |Ax| \geq \frac{1}{4}x_{n_1}^*. \]
4.4.1 Probability Estimate for $\Omega^*_2$.

The proof of Lemma 4.15 will be broken into the following two lemmas.

**Lemma 4.17.** The following probability estimate holds:

$$\mathbb{P}\{\Omega^c_W | \Omega_1\} \leq \exp(-2pn).$$

**Proof.** The proof is essentially the same as that of Lemma 4.11. We will only outline the proof.

We condition on a subevent $\Omega^*_b(\mathcal{A})$ of $\Omega_1$. Furthermore, we condition on $\mathcal{C} = \mathcal{C}_0$, $\mathcal{R} = \mathcal{R}_0$. It is enough to show

$$\mathbb{P}\{\Omega^c_W | \Omega^*_b(\mathcal{A}), \mathcal{R} = \mathcal{R}_0, \mathcal{C} = \mathcal{C}_0\} \leq \exp(-2pn).$$

As a subevent of $\Omega_1$, $|\mathcal{C}_0|$ is either 0 or 1 by Proposition 4.14. Here we pick $\mathcal{C}_0$ to be a subset of size 1, since there is nothing needed to prove when $\mathcal{C}_0 = \emptyset$.

Next, we used $|\mathcal{R}| \leq 3pn|\mathcal{C}| \leq 3pn$ and the same estimate in Lemma 4.11 to obtain the inequality (25) from the proof of Lemma 4.11

$$\mathbb{P}\{|\mathcal{R} \cap \text{supp}(C_j(A))| \geq l | \Omega^*_b(\mathcal{A}), \mathcal{R} = \mathcal{R}_0, \mathcal{C} = \mathcal{C}_0\} \leq n^{-0.9l}.$$

Now, taking $l = \frac{1}{2}t$, and the union bound we get

$$\mathbb{P}\{\Omega_{2\mathcal{R}} | \Omega^*_b(\mathcal{A}), \mathcal{R} = \mathcal{R}_0, \mathcal{C} = \mathcal{C}_0\} \leq n^{1-0.9t} \leq \exp(-2pn).$$

\[\square\]

**Lemma 4.18.**

$$\mathbb{P}\{\Omega^*_b | \Omega_1 \cap \Omega_{2\mathcal{R}}\} \leq \exp(-2pn).$$

**Proof.** The proof is the same as the proof of Lemma 4.12 except some technical details. We will only outline the proof and the specific difference.

We condition on $\mathcal{C} = \mathcal{C}_0$, $\mathcal{R} = \mathcal{R}_0$, and $\Omega^*_b(\mathcal{D})$ which is a sub-event of $\Omega_1 \cap \Omega_{2\mathcal{R}}$.

Now we fix $n_1$ and $n_2$ (which represents the size of $J_1$ and $J_2$) where $n_1$ satisfies the assumption in the lemma and $n_2 \leq \max\{\lfloor\frac{1}{480e}\rfloor, \frac{pn}{3\log^2(pn)}n_1\}$.

First,

$$|\{J_1 \subset J_2 \subset [n] \setminus \mathcal{C} : |J_1| = n_1 \text{ and } |J_2| = n_2\}| \leq \exp\left(\log\left(\frac{2en}{n_2}\right)\right).$$

For a fixed pair of $J_1 \subset J_2$, by Lemma 4.15 with an appropriate choice of $r > 0$,

$$\mathbb{P}\{I_\mathcal{D}(J_1, J_2) = 0 | \mathcal{C} = \mathcal{C}_0, \mathcal{R} = \mathcal{R}_0, \Omega^*_b(\mathcal{D})\} \leq C_{n_1^{1/2}} \exp\left(-\log\left(\frac{2en}{n_2}\right)\frac{r}{16}\right).$$

**Case 1:** $1 \leq n_1 \leq \sqrt{n}$.

Notice that $|\text{supp}(C_j(\mathcal{M}))| \leq \frac{t}{2}$ for every $j \in [n] \setminus \mathcal{C}$. We could apply $r = \frac{t}{2}$. Also, in this case $n_2 \leq n^{0.6}$, together we have $\log\left(\frac{2en}{n_2} \frac{r}{480e(pn)^2}\right) \geq 1$. Next, using $n_2 \leq 3\frac{pn}{\log^2(pn)}n_1$ we obtain

$$\mathbb{P}\{I_\mathcal{D}(J_1, J_2) = 0 | \mathcal{C} = \mathcal{C}_0, \mathcal{R} = \mathcal{R}_0, \Omega^*_b(\mathcal{D})\} \leq \exp(-\log^2(pn)n_2).$$
Applying an union bound we get
\[
\mathbb{P}\{\exists J_1 \subset J_2 \subset [n] \setminus \mathcal{C} : |J_1| = n_1, |J_2| = n_2 \text{ and } I(J_1, J_2) = 0 | \mathcal{C} = \mathcal{C}_0, \mathcal{R} = \mathcal{R}_0, \Omega_\vec{b}(\mathcal{D})\} \\
\leq \exp\left(-\frac{1}{2}\log^2(pn)n_2\right) \\
\leq \frac{1}{n^2} \exp(-2pn).
\]

Case 2: \(\sqrt{n} \leq n_1 \leq \frac{1}{480\exp}\)

When \(\tau \geq 10\), \(L_{\frac{1}{2}pn} \leq \frac{1}{2}n^{1/2}\).

To see that, by (22) we have
\[
u(\frac{1}{2}pn) \leq 2e\exp(\log(6)\frac{r}{2pn} - \frac{pn}{2} + \log(\frac{n}{2})) \leq \exp(-\frac{1}{1000}pn),
\]
where \(\frac{\log(6)}{2} \leq \frac{9}{10} - \frac{1}{1000}\) and \(\log(n) \leq \frac{1}{10}pn\) are used. Next, by (23), \(L_{\frac{1}{2}pn} \leq \frac{6000}{5} \leq \frac{1}{2}n^{1/2}\).

Thus, when \(n_1 \geq \sqrt{n}\), we could substitute \(r = \frac{1}{2}pn\).

Together with \(n_2 \leq \frac{1}{480\exp}\), we have \(\log\left(\frac{2^n}{n_2 \frac{r}{480\exp}(pn)^n}\right) \geq 1\).

Following the same argument as in Case 1, we obtain
\[
\mathbb{P}\{\exists J_1 \subset J_2 \subset [n] \setminus \mathcal{C} : |J_1| = n_1, |J_2| = n_2 \text{ and } I(J_1, J_2) = 0 | \mathcal{C} = \mathcal{C}_0, \mathcal{R} = \mathcal{R}_0, \Omega_\vec{b}(\mathcal{D})\} \\
\leq \frac{1}{n^2} \exp(-2pn).
\]

By an union bound argument, we obtain the estimate stated in the Lemma. \(\square\)

4.4.2 Proof of Theorem 4.3

Proof of Theorem 4.3. Now we fixed such sample \(A\) and \(x\).

First, let’s assume the case \(n_1 = 1\).

Suppose \(\sigma_x(1) \in \mathcal{C}\). Without lose of generality, we assume \(1 = \sigma(1)\). Observe that \(\mathcal{R} = \{i \in \mathcal{R} : a_{i1} = 1\}\). Since \(A\) is a sample from \(\Omega_0\), we know \(\mathcal{R}(x) \geq 1\). For \(i \in \mathcal{R}(x)\),
\[
|Ax_i| = |\sum_{j \in [n]} a_{ij}x_j| \geq |x_1| - |\sum_{j \neq 1} a_{ij}x_j|
\]
If \(n_1 = n_2\), then
\[
|Ax_i| = |\sum_{j \in [n]} a_{ij}x_j| \geq (1 - \frac{3pn}{24pn}x_{n_1}^* \geq \frac{7}{8}x_{n_1}^*.
\]

Now, suppose \(n_2 = 2\) and we may assume \(\sigma_x(2) = [2]\). In this case,
\[
|Ax_i| = |\sum_{j \in [n]} a_{ij}x_j| \geq \frac{7}{8}x_{n_1}^* - a_{i2}|x_2^*|
\]
If \(a_{i2}|x_2| < \frac{1}{4}x_{n_1}^*\), then we have \(|Ax| > \frac{1}{4}x_{n_1}^*\).

Otherwise, let \(y = P_Cx\). Then, \(y_1^* = |y_2^*|\) and \(2 \notin \mathcal{C}\). Hence,
\[
y_1^* > 6pn y_2^*, \quad y_1^* \geq \frac{1}{4}x_{n_1}^*.
\]
Let $J_1 = J_2 = \{2\}$. Since $A$ is a sample from event $\Omega$, there exists $i \in [n] \setminus \mathcal{R}$ such that $i \in I_\mathcal{D}(J_1, J_2)$. Therefore, we have

$$|(Ax)_i| = |(Ay)_i| \geq y^*_i - \frac{6pn}{3pn} y^*_i \geq \frac{1}{8} x^*_n.$$ 

In the remaining cases, either $n_1 = 1$ or $n_1 \geq 2$. We set $y = P_Cx$. Due to $|C| \leq 1 \leq \frac{n_2}{3}$, we have

$$y^*_{[n_1/2]} > 24pn y^*_{n_2 + 1}, \quad y^*_{[n_1/2]} \geq x^*_n.$$ 

By setting $J_1 = \sigma_y([n_1/2])$ and $J_2 = \sigma_y([n_2])$, there exists $i \in I_\mathcal{D}(J_1, J_2)$. Then,

$$|(Ax)_i| = |(Ay)_i| \geq y^*_{[n_1/2]} - \frac{6pn}{3pn} y^*_n \geq \frac{1}{2} x^*_n.$$ 

\[ \square \]

## 5 $\mathcal{T}_2$, $\mathcal{T}_3$, and $\mathcal{R}$ vectors

In this section, we will prove the following two Theorems:

**Theorem 5.1.** Suppose $R > 0, C_T > 0$ are sufficiently large, $r > 0$ is sufficiently small and $p = \frac{1}{2R}$. Then, for sufficiently large $n$ we have

$$\mathbb{P}\left\{ \exists x \in \mathcal{R} \text{ s.t. } \|Ax\| \leq n^{-\frac{1}{T} - o_n(1)} \|x\| \right\} \leq \exp(-2pn).$$

**Theorem 5.2.** Suppose $R > 0, C_T > 0$ are sufficiently large, and $r > 0$ is sufficiently small. Then, for sufficiently large $n$

$$\mathbb{P}\left\{ \exists x \in \mathcal{T}_2 \text{ s.t. } \|\mathfrak{A}x\| \leq n^{-\frac{1}{T} - o_n(1)} \|x\| \right\} \leq \exp(-2pn),$$

$$\mathbb{P}\left\{ \exists x \in \mathcal{T}_3 \text{ s.t. } \|\mathfrak{A}x\| \leq n^{-\frac{1}{T} - o_n(1)} \|x\| \right\} \leq \exp(-2pn).$$

The arguments here is the same as that in the paper [20]. Due to the definition of $\mathcal{T}$ and $\mathcal{R}$ vectors are different, we include the proof for readers’ convenience.

Both of the Theorems can be proven by a net-argument approach. Briefly speaking, we will find a sufficiently dense net $\mathcal{N}$ of the corresponding set ( $\mathcal{R}, \mathcal{T}_2,$ or $\mathcal{T}_3$ ) with relatively small cardinality. And prove that for $x \in \mathcal{N}$, $\|Ax\|$ is small with probability $p$. Then, with probability at most $p|\mathcal{N}|$, there exists $y$ in the corresponding set ( $\mathcal{R}, \mathcal{T}_2,$ or $\mathcal{T}_3$ ) such that $\|Ay\|$ is small. Certainly, the net $\mathcal{N}$ needs to satisfy one crucial property: For each $x$ in the corresponding set ( $\mathcal{R}, \mathcal{T}_2,$ or $\mathcal{T}_3$ ), there exists $y \in \mathcal{N}$ such that $\|A(x - y)\|$ is very small.

To achieve that, we need the following:

Let $\Omega_{\text{norm}}$ be the event that $\|\mathfrak{A} - E\mathfrak{A}\| \leq C_{\text{norm}} \sqrt{pn}$ and $\|\mathfrak{A}\| \leq C_{\text{norm}} \sqrt{pn} + pn$ where $C_{\text{norm}} > 0$ satisfies

$$\mathbb{P}\{\Omega_{\text{norm}}\} \leq \exp(-3pn).$$

By Lemma 2.1, we can choose $C_{\text{norm}}$ to be an universal constant.
Also, we want $\mathcal{N}$ to be a net in the corresponding norm: for $x \in \mathbb{R}^n$,

$$||x|| := ||P_e x|| + \sqrt{p_n} ||P_{e^\perp} x||,$$

where $e \in S^{n-1}$ is the unique vector such that every component equals to $\frac{1}{\sqrt{n}}$, $P_e$ is the orthogonal projection to $e^\perp$, and $P_{e^\perp}$ orthogonal projection to the span of $e$.

Instead of using the standard Euclidean norm, the new norm is defined due to $||Ae||$ is much larger than $||Av||$ for $v \in S^{n-1} \cap e^\perp$.

This section breaks into three parts.

The first part is about the net construction and the estimate of its cardinality. For convenience, we will define

$$\mathcal{T}_i' = \{ x \in \mathcal{T}_i : x^*_n = 1 \} = \left\{ \frac{x}{x^*_n} : x \in \mathcal{T}_i \right\}$$

for $i \in \{2, 3\}$.

The second part is about probability estimate of $||Ax|| > 0$ for $x$ in $\mathcal{R}$, $\mathcal{T}_2'$, and $\mathcal{T}_3'$. Roughly speaking, since $x \notin \mathcal{T}_1$, it is not too sparse to apply Rogozin's Theorem 2.2.

The last part will be the proof of Theorem 5.1 and Theorem 5.2.

### 5.1 Net Construction and Cardinality Estimate

Recall that for $n_s \leq k \leq \frac{n}{\log^2(pn)}$, let $B = [k, n]$. And

$$\mathcal{R}_{k}^1 := \left\{ x \in (\mathcal{Y}(r) \cap AC(r, \rho)) \setminus \mathcal{T} : x^*_n = 1, \frac{||x_{\sigma_2}(B)||}{||x_{\sigma_1}(B)||} \geq C_0 \frac{n}{r} \right\}$$

$$\mathcal{R}_{k}^2 := \left\{ x \in \mathcal{Y}(r) \setminus \mathcal{T} : x^*_n = 1, \frac{||x_{\sigma_2}(B)||}{||x_{\sigma_1}(B)||} \geq C_0 \frac{n}{r} \right\}$$

$$AC(r, \rho) := \left\{ x \in \mathbb{R}^n : \exists \lambda \in \mathbb{R} \text{ s.t. } |\lambda| = x^*_n \text{ and } |\{ i \in [n] : |x_i - \lambda| < \rho \lambda \} | > n - |\mathcal{R}_{k}^1| \right\}.$$

Let $\lambda_1 = \frac{1}{\sqrt{2}}$, and $\lambda_t = 3\lambda_{t-1}$ for $t \in [m-1]$ where $m$ is chosen so that $\lambda_m - 1 < C_2^2 p n < 3\lambda_m - 1$. Let $\lambda_m = C_2^2 p n$. For a simple bound of $m$, we have

$$m \leq 2 \log(pn). \quad (27)$$

For $s \in \{2\}$, we define

$$\mathcal{R}_{k, s}^1 := \left\{ x \in \mathcal{R}_{k}^1 : \lambda_t \sqrt{n} \leq ||x_{\sigma_2}(B)|| \leq \lambda_{t+1} \sqrt{n} \right\}.$$

In this section, we will prove the following lemma:

**Lemma 5.3.** Let $R \geq 40$ be a (large) constant. Then there exists $r_0 > 0$ depending on $R$ with the following property. Let $0 < r \leq r_0$, $0 < \rho \leq \frac{1}{\sqrt{2}}$. Let $s \in \{1, 2\}$, $n_s < k \leq \frac{n}{\log^2(pn)}$, $t \leq m$, and $40 \lambda_t \sqrt{n} \leq \epsilon \leq \lambda_t \sqrt{n}$ where $\lambda_t$ and $m$ are defined according to relation above. Then there exists an $\epsilon$-net $\mathcal{N}^s_{k, t} \subset \mathcal{R}_{s, t}^1$ with respect to $||\cdot||$ of cardinality at most $(\frac{\epsilon}{\lambda_t})^{3m n}$.

**Lemma 5.4.** There exists an $\frac{\sqrt{n}}{C_2 \sqrt{p n}}$-net $\mathcal{N}_{k, t}$ in $||\cdot||$-norm of $\mathcal{T}_i'$ for $i = 2, 3$ such that $|\mathcal{N}_{k, t}| \leq \exp(2 \log(pn)n_{s+i-1})$. 

---

34
We will postpone it to the end of this section. We begin with a few propositions that are required to prove the this lemma.

**Proposition 5.5.** Let \( \| \cdot \|_K \) be a norm on \( \mathbb{R}^n \). (i.e. \( l_2 \) and \( l_\infty \) norm) Consider the set
\[
\{ x \in \mathbb{R}^n : \| x \|_K < a, |\text{supp}(x)| \leq l \}
\]
where \( l < n \) are positive integers and \( a > 0 \). For \( \epsilon > 0 \), there exists an \( \epsilon \)-net \( \mathcal{N}(l, a, \epsilon)_K \) such that
\[
\mathcal{N}(l, a, \epsilon)_K \leq \left( \frac{3a \epsilon \ln n}{\epsilon l} \right)^l \quad \epsilon < a
\]
\[
\mathcal{N}(l, a, \epsilon)_K \leq 1 \quad \epsilon \geq a.
\]

In particular, in either cases we always have
\[
\mathcal{N}(l, a, \epsilon)_K \leq \left( \max\{a, 1\} \frac{3en}{l} \right)^l.
\]

**Proof.** Notice that for \( \epsilon > 0 \), we could trivially set \( \mathcal{N}(l, a, \epsilon) = \{ \emptyset \} \).

Now let \( 0 < \epsilon < a \). Let \( I \subset [n] \) be a subset with \( |I| = l \). By the standard volumetric argument, there exists an \( \epsilon \)-net \( \mathcal{N}_I \) of
\[
\{ x \in \mathbb{R}^n : \| x \|_K \leq a \}
\]
with size bounded by \( (1 + \frac{2a}{\epsilon})^l \leq \left( \frac{3a}{\epsilon} \right)^l \). There are \( \binom{n}{l} \) subsets of \( [n] \) with size \( l \), which is bounded by \( \left( \frac{3a}{\epsilon} \right)^l \).

The union of \( \mathcal{N}_I \) for \( I \subset [n] \) with \( |I| = l \) will be \( \mathcal{N}(l, a, \epsilon) \) and its size will be bounded by \( \left( \frac{3a en}{\epsilon l} \right)^l \). \( \square \)

Next, we will partition the support of \( x \) according to \( \sigma_x \) for \( x \in \mathcal{R}_{kl}^s \). The net we will construct for Lemma 5.3 relies on this partition.

Fix an integer \( k \) such that \( n_s \leq k \leq \frac{n}{\log^2(\rho n)} \). For \( x \in \mathbb{R}^n \), let
\[
B_1(x) = \sigma_x([n_s])
\]
\[
B_{11}(x) = \sigma_x([n_1])
\]
\[
B_{1j}(x) = \sigma_x([n_j \setminus [n_{j-1}]) \quad 2 \leq j \leq s
\]
\[
B_2(x) = \sigma_x([k \setminus [n_s])
\]
\[
B_3(x) = \sigma_x([n_{s+2} \setminus [k])
\]

To deal with \( \mathcal{R}_{kl}^1 \), let \( B'_0(x) = \{ i \in [n] : |\lambda_x - x_i| \leq \rho \} \) where \( \lambda_x \) is the \( \lambda \) appeared in \( \mathcal{AC}(\tau, \rho) \) and define \( B_0(x) = B'_0 \setminus (B_1(x) \cup B_2(x) \cup B_3(x)) \) and \( B_4(x) = [n] \setminus B_0(x) \).

For \( \mathcal{R}_{kl}^2 \), we will skip \( B_0 \) and simply set \( B_4(x) = [n] \setminus (B_1(x) \cup B_2(x) \cup B_3(x)) \).

For \( a > 0 \),
\[
\mathcal{R}(a) = \{ P_{B_4(x)} : x \notin T_1 \text{ and } x_{n_s} = a \}.
\]

**Proposition 5.6.** For \( \epsilon > 0, a > 0 \), there exists an \( \epsilon \)-net \( \mathcal{N}_1(a, \epsilon) \) in \( l_\infty \) norm of the set \( \mathcal{R}(a) \) with cardinality bounded by
\[
\exp \left( 2 \log \left( \frac{a}{\epsilon (\rho n)^3} \right) n_s \right) \quad \text{for } \epsilon < \kappa \rho na,
\]
\[
\exp \left( 2 \log \left( ((\rho n)^3) n_s \right) \right) \quad \text{for } \kappa \rho na \leq \epsilon < \left( \kappa \rho n \right)^s a,
\]
\[
1 \quad \text{for } \epsilon > \left( \kappa \rho n \right)^s a.
\]
In particular, we always have

\[ N_1(a, \epsilon) \leq \exp \left( 2 \log \left( \max \left\{ \frac{a}{\epsilon}, 1 \right\} \right) n_s \right) \]

Proof. Observe that for \( x \in \mathcal{R}(a) \), for \( j \in [s] \),

\[ \| P_{B_1(x)} \|_\infty \leq (\kappa pn)^{s-j+1} a, \quad \text{and} \quad |\text{supp}(P_{B_1(x)})| \leq n_j - n_{j-1}, \]

where the first inequality is due to \( x^{|n_{s-j}|} \leq (\kappa pn)^j a \).

Let \( N_{11} = \mathcal{N}(n_1, (\kappa pn)^{s-j+1} a, \epsilon) \) be the net described in Proposition 5.5. For \( 2 \leq j \leq s \), let \( N_j = \mathcal{N}(n_j - n_{j-1}, (\kappa pn)^{s-j+1} a, \epsilon) \).

For \( x \in \mathcal{R}(a) \), there exists \( y_j \in N_j \) for \( j \in [s] \) such that \( \text{supp}(y_j) \subset B_j(x) \) and \( \| y_j - P_{B_1(x)} \|_\infty \leq \epsilon \).

Due to \( \{B_j(x)\}_{j \in [s]} \) are disjoints, we have

\[ \left\| P_{B_1(x)} - \sum_{j \in [s]} y_j \right\|_\infty = \max_{j \in [s]} \left\| P_{B_1(x)} - y_j \right\|_\infty \leq \epsilon. \]

Therefore, the net \( \mathcal{N}_1 = \mathcal{N}_{11} \times \mathcal{N}_{12} \times \ldots \mathcal{N}_{1s} \) will be the \( \epsilon \)-net for \( \mathcal{R}_a \).

First of all, if \( \epsilon \geq (\kappa pn)^s a = x_1^* \), then \( |\mathcal{N}_1| = 1 \).

Now, we assume that \( \epsilon < (\kappa pn)^j a \) and let \( j_0 \) be the largest integer in \([s]\) such that \( \epsilon \geq (\kappa pn)^{s-j_0+1} a = x_1^* \).

Then, for \( j_0 < j \leq s \), \( |\mathcal{N}_j| = 1 = \epsilon^0 \). And for \( 1 \leq j \leq j_0 \),

\[ |\mathcal{N}_j| \leq \exp \left( \log (\frac{3ea(\kappa pn)^{s-j+1} n_j}{\epsilon n_j}) n_j \right) = \exp(N_j). \]

Next, we want to show that \( \frac{N_{j-1}}{N_j} \leq \frac{1}{2} \) for \( 2 \leq j \leq j_0 \). If that is true, then

\[ |\mathcal{N}_1| \leq \exp \left( \sum_{j=1}^{j_0} N_j \right) \leq \exp(2N_{j_0}). \]

Now let us fix such \( j \) if it exists. (i.e. \( j_0 > 1 \).)

\[ N_{j-1} = \log \left( \frac{3ea(\kappa pn)^{s-j+1} n_j}{\epsilon n_j} \right) n_{j-1}. \]

By \( \frac{(\kappa pn)^{s-j+1} a}{\epsilon} \geq 1 \) and \( n_j \leq n_s \leq \frac{1}{480e^3} \),

\[ \frac{3ea(\kappa pn)^{s-j+1} n_j}{\epsilon n_j} \geq pn. \]

Next, by \( \frac{n_j}{n_{j-1}} \leq \frac{pn}{\log^3 (pn)} \),

\[ \frac{n_j}{n_{j-1}} \epsilon (\kappa pn) \leq (pn)^3. \]

Therefore, we conclude that

\[ N_{j-1} \leq 4 \log \left( \frac{3ea(\kappa pn)^{s-j+1} n_j}{\epsilon n_j} \right) n_{j-1} \leq \frac{1}{2} N_j. \]
where we used \( \frac{n_j}{n_{j-1}} \leq \frac{pn}{\log^3(pn)} \) again.

The cardinality of \( \mathcal{N}_1 \) is bounded by \( \exp(2N_{j_0}) \). If \( \epsilon < ak\rho \), then \( j_0 = s \) and

\[
|\mathcal{N}_1| \leq \exp(2 \log(\frac{3eak(pn)n}{\epsilon n_s})) n_s \leq \exp \left( 2 \log \left( \frac{a(pn)^3}{\epsilon} \right) n_s \right)
\]

For \( \epsilon > (\kappa pn)a \), then

\[
\frac{(\kappa pn)^{s-j_0+1}a}{\epsilon} \leq \kappa pn
\]

and

\[
|\mathcal{N}_1| \leq \exp(2 \log(\frac{3ek(pn)n}{n_{j_0}})) \leq \exp(6 \log(pn) n_{j_0})
\]

where we used \( n_j \leq n_s \leq \frac{1}{480\epsilon p} \) for the second inequality.

\[\Box\]

**Proposition 5.7.** For \( n_{s_0} \leq k \leq \frac{n}{\log^3(pn)} \), there exists a \( \rho \sqrt{n} \)-net \( \mathcal{N}_4 \) in \( l_2 \) norm for \( \{PB_{0}(x) : x \in \mathcal{R}_k^1 \} \) with support bounded by \( \exp(2 \log(\frac{\epsilon}{\rho}) \sqrt{2n}) \).

**Proof.** For \( \mathcal{R}_k^1 \), \( |B_{0}(x)'| > n - \sqrt{n} \). Consider the collection \( \{I \subset [n] : |I| \geq n - \sqrt{n} \} \). Its cardinality is \( \sum_{j=0}^n (\binom{n}{j}) \leq \exp(\log(\sqrt{n}) + \log(\frac{\epsilon}{\rho}) \sqrt{n}) \). Let \( \mathcal{N}_4 = \{ \pm P_I(\sqrt{n}) : I \subset [n] \text{ and } |I| \geq n - \sqrt{n} \} \). Then, \( |\mathcal{N}_4| \leq \exp(2 \log(\frac{\epsilon}{\rho}) \sqrt{n}) \). Also, for \( x \in \mathcal{R}_k^1 \), the vector \( y = P_{B_{0}(x)\sqrt{n}} \in \mathcal{N}_4 \) satisfies \( \|P_{B_{0}(x)}x - y\|_2 \leq \rho \sqrt{n} \). \[\Box\]

**Proof of Lemma 5.3** For \( x \in \mathcal{R}_k^s \), \( \|P_{B_3(k,x)}x\| \leq \lambda_{t+1} \sqrt{n} \), we will use different nets to approximate \( P_{B_i}(x) \) for different \( i \).

| index \( i \) | net | \( \log(|\mathcal{N}_i|) \) | \( l_2 \) norm distance |
|---|---|---|---|
| \( i = 1 \) | \( \mathcal{N}_1 := \mathcal{N}_1(C_T^2pn, \epsilon_1) \) | \( 2 \log \left( \max \left( \frac{C_T^2pn}{\epsilon_1}, 1 \right) (pn)^3 \right) n_s \) | \( \sqrt{n} \epsilon_1 \) |
| \( i = 2 \) | \( \mathcal{N}_2 := \mathcal{N}(k - n_s, C_T^2pn, \epsilon_2) \) | \( \log \left( \max \left( \frac{C_T^2pn}{\epsilon_2}, 1 \right) \frac{3en}{k} \right) k \) | \( \sqrt{k} \epsilon_2 \) |
| \( i = 3 \) | \( \mathcal{N}_3 := \mathcal{N}(\sqrt{n}k - k, \lambda_{t+1} \sqrt{n}, \epsilon_3)_2 \) | \( \log \left( \frac{\lambda_{t+1} \sqrt{n}}{\epsilon_3} \right) \sqrt{n}k \) | \( \epsilon_3 \) |
| \( i = 0 \) | \( \mathcal{N}_0 := \{ \pm P_I(\sqrt{n}) : I \subset [n] \text{ and } |I| \geq n - \sqrt{n} \} \) | \( 2 \log(\frac{\epsilon}{\rho}) \sqrt{n} \) | \( \rho \sqrt{n} \) |

Let \( \mathcal{N} = \mathcal{N}_1 \times \mathcal{N}_2 \times \mathcal{N}_3 \times \mathcal{N}_4 \). For \( x \in \mathcal{R}_k^{s_1} \), there exists \( y_i \in \mathcal{N}_i \) for \( i = 0, 1, 2, 3 \) such that

\[
\left\| x - \sum_{i=0}^3 y_i \right\|^2 \leq \sum_{i=0}^3 \left\| P_{B_i}(x) - y_i \right\|^2 \leq \rho^2 n + n_s \epsilon_1^2 + k \epsilon_2^2 + \epsilon_3^2 + \sqrt{n} \quad (28)
\]

Let \( \epsilon_3 = \frac{1}{3} \epsilon, \epsilon_1 = \frac{1}{\sqrt{n}} \epsilon_3, \) and \( \epsilon_2 = \frac{1}{\sqrt{k}} \epsilon_3 \). Then, \( (28) \leq (\rho^2 + \sqrt{n}) n + \frac{1}{3} \epsilon^2 \leq \frac{2}{3} \epsilon^2 \) when \( \epsilon > 0 \) is sufficiently small.

Next, we will bound the net size. First, from the choice of \( \epsilon \) and \( \lambda_{t+1} \geq \frac{1}{\sqrt{2}}, \epsilon_1, \epsilon_2 \geq \frac{1}{\sqrt{k}} \). Together with \( n_s \leq \frac{k}{\log^3(pn)} \),
\[
\log(|\mathcal{N}_1|) \leq 8 \log(pn)n_s,
\]
\[
\log(|\mathcal{N}_2|) \leq 3 \log(pn)k
\]
\[
\log(|\mathcal{N}_3|) \leq \log\left(\frac{9R}{40r}\right)rn \leq 2 \log\left(\frac{1}{r}\right)rn
\]

Therefore, we have
\[
\log(|\mathcal{N}|) \leq \left(\frac{1}{r}\right)^{\frac{3}{2}rn}.
\]

For \(\mathcal{R}_{kt}^2\), we set \(\mathcal{N} = \mathcal{N}_1 \times \mathcal{N}_2 \times \mathcal{N}_3\) with the same choice of \(\epsilon_1, \epsilon_2, \text{and } \epsilon_3\). Similarly, for \(x \in \mathcal{R}_{kt}^2\) there exists \(y \in \mathcal{N}\) satisfying
\[
\|x - y\| \leq \frac{1}{3} \epsilon^2 + n \leq \frac{2}{3} \epsilon^2
\]
where the last inequality holds when \(\lambda_t > \frac{2}{3}\) and \(r\) is sufficiently small. And this is sufficient since for \(\lambda_t < \frac{2}{3}\), \(\mathcal{R}_{kt}^2\) is an empty set due to the definition of \(\mathcal{R}_{kt}^2\).

Therefore, we have constructed a \(\sqrt{\frac{2}{3} \epsilon}\)-net for the corresponding \(\mathcal{R}_{kt}^2\) in the \(l_2\)-norm.

It remains to give the \(||\cdot||\) distance estimate using the same net.

By Lemma 5.3, we knew that for \(x \in \mathcal{R}_k\), \(x_1 \leq n^{1+o(1)}\) and thus, \(|\langle x, e \rangle| \leq \frac{1}{\sqrt{n}} \sum_{i \in [n]} x_i^* \leq n^{1.6}\).

Let \(\mathcal{N}_e\) be an \(0.1e\)-net of the set \(\{te : |t| \leq n^{1.6}\}\). Then, \(|\mathcal{N}_e| \leq n^2\).

Let \(\mathcal{N}_{kt} = \{Pe y + v : y \in \mathcal{N}, v \in \mathcal{N}_e\}\). Then, \(|\mathcal{N}_{kt}| \leq \left(\frac{2}{3}\right)^{3n}\).

For \(x \in \mathcal{R}_{kt}^2\), let \(y \in \mathcal{N}\) and \(v \in \mathcal{N}_e\) such that \(||x - y\|| \leq \sqrt{\frac{2}{3} \epsilon}\) and \(||P_e^\perp v\|| \leq \frac{0.1\epsilon}{pn}||P_e x - v|| \leq \epsilon\).

Then,
\[
||x - Pe y + v\in\mathcal{N}_{kt}|| = ||P_e x - P_e y|| + pn||P_e^\perp x - v|| \leq ||x - y|| + pn||P_e^\perp x - v|| \leq \epsilon.
\]

\(\Box\)

**Proof of Lemma 5.4** The proof is almost identical to that of Lemma 5.3. So we just show the case for \(\mathcal{T}_3\).

Let \(k = n_{s+1}\) and we keep the same definition of \(B_1(x), B_2(x), \text{and } B_3(x)\).

For \(x \in \mathcal{T}_3\), we have \(x_{n_s} \leq C_T \sqrt{pn}, x_{n_{s+1}} \leq 1, \text{and } x_{n_{s+2}} \leq \frac{1}{C_T \sqrt{pn}}\). Thus, we use the following net:

| index \(i\) | net | \(\log(|\mathcal{N}_i|)\) | \(\forall x \in \mathcal{T}_3 \exists y \in \mathcal{N} \text{ s.t. } ||p_{B_i}(x) - y||_2 \leq \) |
|---|---|---|---|
| \(i = 1\) | \(\mathcal{N}_1 := \mathcal{N}_1(C_T \sqrt{pn}, \frac{1}{2 C_T \sqrt{pn}})\) | \(10 \log(pn)n_s\) | \(\sqrt{n_s} \frac{1}{2 C_T \sqrt{pn}}\) |
| \(i = 2\) | \(\mathcal{N}_2 := \mathcal{N}(k - n_s, C_T \sqrt{pn}, \frac{1}{2 C_T \sqrt{pn}})\) | \(3 \log(pn)k\) | \(\frac{1}{2 C_T \sqrt{pn}}\) |
| \(i = 3\) | \(\mathcal{N}_3 := \mathcal{N}(\lfloor r n \rfloor - k, 1, \frac{1}{C_T \sqrt{pn}})\) | \(\log(3eC_T \sqrt{pn})rn\) | \(\sqrt{R_{s_0} + 2 C_T \sqrt{pn}}\) |

Then, \(\mathcal{N} = \mathcal{N}_1 \times \mathcal{N}_2 \times \mathcal{N}_3\) will be the net which approximates \(\mathcal{T}_3\) well in the \(l_2\) norm. To pass it to the triple norm \(||\cdot||\). Let
\[
\mathcal{N}_{T_3} := \{Pe y + v : y \in \mathcal{N} \text{ and } v \in \mathcal{N}_e\}
\]
where \(\mathcal{N}_e\) is the net appeared in the proof of Lemma 5.3.

\(\Box\)
5.2 Tail bound for individual probability

In this section, we will give a probability estimate of $\|Ax\| > 0$ for $x \in \mathcal{R}_k$, $x \in \mathcal{T}_2^f$, and $x \in \mathcal{T}_3^f$, respectively.

**Proposition 5.8.** For $x \in \mathcal{R}^s_{kt}$, let $I = \sigma_x([k, n])$. We have

$$
\mathbb{P} \left( \|Ax\| \leq \frac{1}{2\sqrt{10}} C_{Rgz} \sqrt{pn} \lambda_1 \sqrt{n} \right) \leq \exp \left( -\frac{n}{6} \right).
$$

**Proof.** By definition of $\mathcal{R}^s_{kt}$, $\|x\|_\infty \geq \frac{2C_{Rgz}}{\sqrt{p}}$.

Applying (5) we get

$$
\mathcal{Q} \left( \frac{(Ax)_i \sqrt{p}}{2C_{Rgz}} \right) \leq \frac{1}{2}.
$$

Due to $\{(Ax)_i\}_{i \in [n]}$ are jointly independent. By (8), we have

$$
\mathbb{P} \left( \left\{ i \in [n] : |(Ax)_i| > \frac{\|x\|_\infty \sqrt{p}}{2C_{Rgz}} \right\} \right) \leq \frac{n}{10} \leq \exp \left( -\frac{n}{6} \right).
$$

If the complement of the event in the above inequality holds, then

$$
\|Ax\| \geq \sum_{i \in [n] \text{ s.t. } |(Ax)_i| > \frac{\|x\|_\infty \sqrt{p}}{2C_{Rgz}}} (Ax)_i^2 \geq \frac{n}{10} \frac{p}{2C_{Rgz}} \|x\|_2^2.
$$

Therefore,

$$
\mathbb{P} \left( \|Ax\| \leq \frac{1}{2\sqrt{10}} C_{Rgz} \sqrt{pn} \|x\| \right) \leq \exp \left( -\frac{n}{6} \right).
$$

Finally, the statement of the Proposition follows due to $\|x\| \geq \lambda_1 \sqrt{n}$ by the definition of $\mathcal{R}^s_{kt}$. 

**Lemma 5.9.** Let $A$ be a $n \times n$ random Bernoulli matrices with parameter $p$. Assuming that $n$ is sufficiently large. For a vector $x \in \mathbb{R}^n$ satisfying

$$
x_{m_1}^* = 3a \quad \text{and} \quad x_{n_1}^* > 3x_{n-m_1}^*.
$$

where $m_0 \leq m_1 \leq \frac{n}{2}$ and $a > 0$. We have

$$
\mathbb{P} \left\{ \|Ax\| < \sqrt{\frac{1}{50} qn} \right\} \leq \exp \left( -\frac{1}{40} qn \right)
$$

where $q = 2m_0p(1-p)^{2m_0-1}$. Furthermore, if $\frac{m_1}{m_0} q$ is greater than an universal constant, we have

$$
\mathbb{P} \left\{ \|Ax\| < \sqrt{\frac{n}{4q}} \right\} \leq 2 \exp \left( -\frac{1}{12} \log(\frac{m_1}{m_0} q)n \right).
$$

Remark: We will apply this lemma with $m_0 = n_s = \frac{1}{480p}$ and $m_1 = m_0$ or $m_1 = n_{s+1}$. In either cases, $q \geq m_1 p$. As a Corollary, we have
Corollary 5.9.1. We have the following estimate,

\[
\begin{align*}
for \ x \in T_2' \quad & \mathbb{P} \left\{ \|Ax\| < \sqrt{\frac{5.9.1}{n}} \right\} \leq \exp(-c_{5.9.1}n), \\
for \ x \in T_3' \quad & \mathbb{P} \left\{ \|Ax\| < \frac{1}{6} \sqrt{n} \right\} \leq \exp\left( -\frac{1}{24} \log(pn)n \right).
\end{align*}
\]

Proof. We will apply Lemma 5.9 with \( \mathbb{A} \). Suppose \( m_0 \geq n_s \), due to \( n_s = \lfloor \frac{1}{480 np} \rfloor \),

\[
(1 - p)^{2m_0 - 1} \leq (1 - p)^{m_0} \leq \exp(-pm_0) \leq \frac{1}{2}
\]

and hence,

\[ q \geq m_0 p. \]

For \( x \notin T_2' \), we apply Lemma 5.9 with \( m_0 = m_1 = n_s, \) and \( a = 1 \):

\[ \mathbb{P} \left\{ \|Ax\| < \sqrt{\frac{5.9.1}{n}} \right\} \leq \exp(-c_{5.9.1}n) \]

for a sufficiently small universal constant \( c_{5.9.1} > 0 \).

For \( x \notin T_3' \), we apply Lemma 5.9 with \( m_0 = n_s, \) \( m_1 = n_{s+1}, \) and \( a = 1 \):

\[ \mathbb{P} \left\{ \|Ax\| < \sqrt{\frac{n}{4}} \right\} \leq 2 \exp\left( -\frac{1}{12} \log_{\frac{1}{2}}(mpn) \right) \leq \exp\left( -\frac{1}{24} \log(pn)n \right). \]

Before we move on to the proof. We will set up some notations. Let \( h = \left\lfloor \frac{m_1}{m_0} \right\rfloor \). We define the corresponding sets:

\[
\begin{align*}
J'_t &= \sigma_x((tm_0)\setminus[(t-1)m_0]) & t & \in \left[ h \right] \\
J^*_t &= \sigma_x([n-(t-1)m_0]\setminus[n-tm_0]) & t & \in \left[ h \right] \\
J_t &= J'_t \cup J^*_t \\
J_0 &= [n] \setminus \left( \bigcup_{t \in [h]} J_t \right) \\
J &= (J_1, \ldots, J_h).
\end{align*}
\]

For \( J \subset [n] \), we define

\[ I(J) = \{ i \in [n] : \exists j_0 \text{ s.t. } a_{ij_0} = 1 \text{ and } a_{ij} = 0 \text{ for } j \in J \setminus \{ j_0 \} \}. \]

Next, let \( \mathcal{I} \) be a collection of subsets of \([n]\) : \( \mathcal{I} = (I_1, \ldots, I_s) \) where \( I_i \subset [n] \). Let \( M \) be a \([n] \times J_0\) matrix with 0,1 entries. Specifically, \( M \) is a \([n] \times |J_0|\) matrix whose columns are indexed by \( J_0 \) and rows are indexed by \([n]\).

Let \( \Omega_{J,\mathcal{I},M} \) be the event of \( A \) that \( 1. I(J_t) = I_t \) for \( t \in [h] \), and \( 2. A_{J_0} = M \), where \( A_{J_0} \) is the submatrix of \( A \) with columns \( J_0 \).

For \( i \in [n] \) and \( J \subset [n] \), let \( A_{i,J} \) denotes the \( 1 \times |J| \) submatrix of \( A \) with row \( i \) and columns \( J \).

Now we condition on \( \Omega_{J,\mathcal{I},M} \), then \( \{ A_{i,J_t} \}_{t \in [h], i \in [n]} \) are jointly independent but not necessary i.i.d.

Fix \( i \in [n] \), we have

\[(Ax)_i = \sum_{t=0}^{h} \sum_{j \in J_t} a_{ij} x_j := \sum_{t=0}^{h} \xi_t.\]
Proposition 5.10. Condition on $\Omega J, I, M$ and fix $i \in [n]$ and $t \in [h]$ such that $i \in I(J_t)$. Let $\xi_t := \sum_{j \in J_t} a_{ij} x_j$. Then, $\mathcal{D}(\xi_t, 2a) \leq \frac{1}{2}$.

Proof. Now fix $i \in I(J_t)$, the $A_{ij}$ contains exactly one none-zero entry: There exists $j_0$ which is uniformly chosen among $J_t$ such that $a_{ij_0} = 1$ and $a_{ij} = 0$ for $j \in J \setminus \{j_0\}$. Notice that $\xi_t = x_{j_0}$.

First, $\mathbb{P}\{j_0 \in J_t^1\} = \mathbb{P}\{j_0 \in J_t^r\} = \frac{1}{2}$. Secondly, for $j_1 \in J_t^1$, $j_r \in J_t^r$, we have $|x_{j_1}| > 3|x_{j_r}| \geq a$ by the definition of $J_t$. Together we conclude that $\mathcal{D}(x_{j_0}, a) \leq \frac{1}{2}$. \hspace{1cm} $\square$

Proof of Lemma 5.9. Consider the set

$$S_i := \{t \in [h] : i \in I(J_t)\}.$$ 

If we condition on $\Omega J, I, M$, the set $S_i$ is fixed and determined by $I$. In the case that $S_i$ is non-empty, let $t \in S_i$ and we have $\mathcal{D}((Ax)_i, a) \leq \mathcal{D}(\xi_t, a) \leq \frac{1}{2}$.

Case 1: $h = 1$ In this case, $J = (J_1)$. For simplicity, let $J := J_1$ and $I := I_1$. The expected size of $I(J)$ is $nq$. Let $O_1$ be the event that $|I(J)| \geq \frac{1}{3} nq$, by (8) we have

$$\mathbb{P}\{O^c_1\} \leq \exp(-\frac{1}{3} nq).$$

(29)

Notice we could partition the event $O_1$ into subevents of the form $\Omega J, I, M$.

Now we condition on an event $\Omega J, I, M \subset O_1$. In this case, $I(J) = I$ is a fixed set.

Since $(Ax)_i \in I(J)$ are jointly independent, we have

$$\mathbb{P}\left\{|\{i \in I(J) : |(Ax)_i| > a\}| \leq \frac{1}{2} |I(J)| \bigg| \Omega J, I, M\right\} \leq \exp(-\frac{1}{3} \cdot \frac{1}{2} |I(J)|) \leq \exp(-\frac{1}{30} nq).$$

by (8). Hence,

$$\mathbb{P}\left\{|\{i \in I(J) : |(Ax)_i| > a\}| \leq \frac{1}{50} nq \bigg| O_1\right\} \leq \exp(-\frac{1}{30} nq).$$

(30)

Let $O_2$ be the event that $|\{i \in I(J) : |(Ax)_i| > a\}| \geq \frac{1}{50} nq$. By (29) and (30),

$$\mathbb{P}\{O^c_2\} \leq \mathbb{P}\{O^c_1\} + \mathbb{P}\{O^c_2 \big| O_1\} \leq \exp(-\frac{1}{40} nq).$$

Furthermore, within the event $O_2$, we have $\|Ax\|^2 \geq \frac{1}{50} nqa^2$.

Case 2: $h > 1$ The expected size of $|S_i|$ is $hq$. Recall that $A_{ij}$ are jointly independent for $t \in [h]$. Thus, applying (8) we get

$$\mathbb{P}\left\{|S_i| < \frac{1}{5} hq\right\} \leq \exp(-\frac{1}{3} hq) := q_2.$$ 

Let $I := \{i \in [n] : |S_i| > \frac{1}{5} hq\}$ and $O_1$ be the events that $|I| \geq \frac{n}{2}$. By (8),

$$\mathbb{P}\{O^c_1\} = \mathbb{P}\left\{|I| \geq \frac{n}{2}\right\} \leq \left(\frac{enq_2}{n/2}\right)^{n/2} \leq \exp(-\frac{1}{12} \log(hq)n).$$

provided that $hq > 6 \log(2e)$.
We could also partition $O_1$ into subevents of the form $\Omega_{J,I,M}$. Now we fix a subevent $\Omega_{J,I,M}$ of $O_1$. By Rogozin’s Theorem (Theorem 2.2), we have
\[
\mathcal{Q}(\sum_{t \in S_i} \xi_t, a) \leq \frac{C}{\sqrt{\frac{1}{2} |S_i|}} \leq \frac{C \sqrt{10}}{\sqrt{hq}}
\]
for $i \in I$. Let
\[
I' := \{ i \in I_1 : |(Ax)_i| > a \}.
\]
Due to independence of the rows (after conditioning on $\Omega_{J,I,M}$), by (6) we have
\[
\mathbb{P}\left( |I'| > \frac{|I|}{2} \mid \Omega_{J,I,M} \right) \leq \exp\left( \log\left( \frac{2eC \sqrt{10}}{\sqrt{hq}} \right) \frac{|I|}{2} \right) \leq \exp\left( -\log(hq) \frac{|I|}{5} \right),
\]
and thus
\[
\mathbb{P}\left( | \{ i \in I_1 : |(Ax)_i| > a \} | \leq \frac{n}{4} \mid O_1 \right) \leq \exp\left( -\log(hq) \frac{n}{10} \right).
\]
Thus, let $O_2$ be the event that $|I'| > \frac{n}{4}$, we have
\[
\mathbb{P}\{ O_2 \} \leq \mathbb{P}\{ O_1^c \} + \mathbb{P}\{ O_2^c \mid O_1 \} \leq 2 \exp\left( -\frac{1}{12} \log(hq)n \right).
\]
When $O_2$ holds, we have $\|Ax\| \geq \sqrt{\frac{n}{4}} a$. 

5.3 Estimate for $T'_2$, $T'_3$ and $R_{kt}^s$

Proof of Theorem 5.1. We know that
\[
\mathcal{R} := \bigcup_{s \in [2]} \bigcup_{n_s \leq k \leq \frac{4}{\log^2(p\lambda)}} \bigcup_{t \in [m]} \mathcal{R}_{kt}^s.
\]
Now we focus on $\mathcal{R}_{kt}^s$ for a triple $(s, k, t)$.

Let $\mathcal{N} := \mathcal{N}_{kt}^s$ be the net described in Lemma 5.3. The cardinality of the net is bounded by $(\frac{\varepsilon}{t})^{3n}$. When $t$ is sufficiently small, we may assume $\mathcal{N} \leq \exp\left( \frac{n}{12} \right)$.

By Proposition 5.8 and the union bound argument, we have
\[
\mathbb{P}\left\{ \exists x \in \mathcal{N} \text{ s.t. } \|Ax\| \leq \frac{CRgz}{2\sqrt{10}} \sqrt{p\lambda t \sqrt{n}} \right\} \leq \exp\left( -\frac{n}{12} \right).
\]
Now we condition on the event that $\|Ax\| \geq \frac{CRgz}{2\sqrt{10}} \sqrt{p\lambda t \sqrt{n}}$ for all $x \in \mathcal{R}_{kt}^s$ and $\Omega_{\text{norm}}$. For $x \in \mathcal{R}_{kt}^s$, $\exists y \in \mathcal{N}$ such that $\|x - y\| < \frac{40}{R} \lambda t \sqrt{n}$. Also,
\[
\|Ax\| \geq \|Ay\| - \|A(x - y)\| \geq \|Ay\| - \|(A - \mathbb{E}Ax)(x - y)\| - \|\mathbb{E}Ax - y\| \geq \frac{CRgz}{2\sqrt{10}} \sqrt{p\lambda t \sqrt{n}} - \frac{40}{R} \lambda t \sqrt{n} C_{\text{norm}} \sqrt{p\lambda t \sqrt{n}} \geq \frac{CRgz}{20} \sqrt{p\lambda t \sqrt{n}}.
\]
and the last inequality holds when $R$ is sufficiently large.
Furthermore, since for any \(x \in \mathcal{R}, x \notin \mathcal{T}\). By Proposition 3.4, \(\|x\| \leq n^{1+o_n(1)}\). And by the definition of \(\mathcal{R}_{kt}^s, \lambda t \sqrt{n} \geq \sqrt{\frac{r}{2}}\), we obtain

\[
\lambda t \sqrt{n} \geq n^{-\frac{1}{2} - o_n(1)}.
\]

and hence,

\[
\mathbb{P}\left\{ \exists x \in \mathcal{R}_{kt}^s \text{ s.t. } \|Ax\| \leq n^{-\frac{1}{2} - o_n(1)} \|x\| \mid \Omega_{\text{norm}} \right\} \leq \exp\left(-\frac{1}{24}n\right).
\]

Finally, the statement of the Theorem follows by applying a union bound arguments over all possible triple \((s, k, t)\) and \(\mathbb{P}(\Omega_{\text{norm}}^c) \leq \exp(3\log(pn)n_{s+2})\). □

**Proof of Theorem 5.2**. The proof for Theorem 5.2 is essentially the same as that of Theorem 5.1. Instead of using Lemma 5.3 and Proposition 5.8 for the net and individual probability estimate, we can replace it with Lemma 5.4 and Corollary 5.9.1. Here we will sketch the proof for the case \(T'_3\) only.

By Lemma 5.4, there exists a \(\sqrt{\frac{2n}{C_T \sqrt{pn}}}\)-net, \(\mathcal{N}\) in \(\|\cdot\|\)-norm for \(T'_3\) whose size is bounded by \(\exp(2 \log(\frac{pn}{n_s+2})\).

By Corollary 5.9.1,

\[
\mathbb{P}\left\{ \exists x \in \mathcal{N} \text{ s.t. } \|Ax\| \leq \frac{1}{6} \sqrt{n} \right\} \leq \exp\left(-\frac{1}{24} \log(pn)n + 2 \log(pn)rn\right) \leq \exp\left(-\frac{1}{48} \log(pn)n\right),
\]

when \(r > 0\) is sufficiently small. Conditioning on the event that \(\|Ax\| \geq \frac{1}{6} \sqrt{n}\) for all \(x \in \mathcal{N}\). For \(x \in T'_3\), there exists \(y \in \mathcal{N}\), such that \(\|x - y\| \leq \frac{\sqrt{2n}}{C_T \sqrt{pn}}\). If we condition on \(\Omega_{\text{norm}}\),

\[
\|Ax\| \geq \frac{1}{6} \sqrt{n} - 2 \frac{C_{\text{norm}}}{C_T} \sqrt{2n} \geq \frac{1}{12} \sqrt{n}
\]

when \(C_T > 100 C_{\text{norm}}\). Together with \(\mathbb{P}\{\Omega_{\text{norm}}^c\} \leq \exp(-3\log(pn)n)\) and the definition of \(T'_3\),

\[
\mathbb{P}\left\{ \exists x \in T'_3 \text{ s.t. } \|Ax\| \leq \frac{1}{6} \sqrt{\|x\|_{n_{s+1}}} \right\} \leq \exp(-2\log(pn)n).
\]

In the end, for \(x \in T_3, x \notin T_1\). By Proposition 3.4, \(\|x\|_{n_{s+1}} \leq n^{1+o_n(1)}\) for \(x \in T_3\). Thus, we finish the proof. □

**6 Proof of Main Theorem**

Let \(\Omega_{RC}\) be the event that no row or column of \(\mathcal{A}\) is \(\vec{0}\). As a Corollary of Theorem 3.7, 4.1, 5.1 and 5.2, we have

**Theorem 6.1.** There exists large constant

\[
R > 1 \quad \kappa > 1
\]

and small constants

\[
r > 0 \quad \rho = \frac{1}{2R} \quad \delta = \frac{1}{3^r}
\]

43
such that the following holds:

Let \( g \) be the growth function defined in Definition 3.5. For sufficiently large \( n \), there exists an event \( \Omega \) with
\[
\mathbb{P}\{\Omega^c\} \leq n \exp(-\frac{1}{10} pn)
\]
so that if we condition on \( \Omega_{RC} \cap \Omega \), for any \( x \in \mathcal{V}(t) \backslash \mathcal{V}(r, g, \delta, \rho) \),
\[
\|Ax\| \geq n^{-1-o_n(1)}\|x\| \quad \text{and} \quad \|A^\top x\| \geq n^{-1-o_n(1)}\|x\|.
\]

Next, we want to combine Theorem 6.1 and 3.1 to get Theorem 1.4. If we used the same sets of parameters, by Proposition 3.6, both \( a_n \) and \( b_n \) is \( n^{1+o_n(1)} \). Furthermore, the event \( E \) from Theorem 3.1 and the event \( \Omega \) from Theorem 6.1 satisfies \( E^c \subset (\Omega_{RC} \cap \Omega)^c \).

Thus, combining the theorem we have the following inequalities: For \( t \geq 0 \),
\[
\mathbb{P}\{s_{\min}(\mathcal{A}) \leq n^{-2+o_n(1)}t\} \leq \mathbb{P}\{E^c\} + \mathbb{P}\{\|Ax\| \leq n^{-2+O_n(1)}t \text{ for some } x \in \mathcal{V}_n \cap \mathcal{E}\}
\]
\[
\leq (1 + o_n(1))\mathbb{P}(\Omega_{RC}^c) + n^{o_n(1)}t,
\]
where we rely on \( \mathbb{P}\{\Omega_{RC}^c\} = O_n(n \exp(-pn)) \) from Lemma 2.5.

By resetting the value \( t \) we get
\[
\mathbb{P}\{s_{\min}(\mathcal{A}) \leq n^{-2+o_n(1)}t\} \leq (1 + o_n(1))\mathbb{P}\{\Omega_{RC}^c\} + t.
\]

Next, by Lemma 2.5 we have
\[
\mathbb{P}\{\Omega_{RC}^c\} = (1 + o_n(1)) \left(1 - (1 - (1 - p)^n)^{2n}\right)
\]
and we finish the proof of Theorem 1.4.

References

[1] A. Basak and M. Rudelson. Sharp transition of the invertibility of the adjacency matrices of random graphs. arXiv:1809.08454.

[2] J. Bourgain, V. Vu, and P. Wood. On the singularity probability of discrete random matrices. J. Funct. Anal., 258(2):559–603, 2010.

[3] J. Kahn, J. Komlós, and E. Szemerédi. On the probability that a random \( \pm \)-matrix is singular. J. Amer. Math. Soc., 8(1):223–240, 1995.

[4] J. Komlós. On the determinant of \((0, 1)\) matrices. Studia Sci. Math. Hungar, 2:7–21, 1967.

[5] E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. The circular law for sparse random regular digraphs. J. European Math. Soc. to appear.

[6] E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. Anti-concentration property for random digraphs and invertibility of their adjacency matrices. C. R. Math. Acad. Sci. Paris, 354(2):121–124, 2016.

[7] E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. Adjacency matrices of random digraphs: singularity and anti-concentration. J. Math. Anal. Appl., 445(2):1447–1491, 2017.

[8] E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. The rank of random regular digraphs of constant degree. J. of Complexity, 48:103–110, 2018.
[9] E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. Structure of eigenvectors of random regular digraphs. *Tran. Amer. Math. Soc.*, 371:8097–8172, 2019.

[10] E. Litvak, A. Lytova, K. Tikhomirov, N. Tomczak-Jaegermann, and P. Youssef. The smallest singular value of a shifted-d-regular random square matrix. *Probab. Theory Related Fields*, 173:1301–1347, 2019.

[11] G. Livshyts. The smallest singular value of heavy-tailed not necessarily i.i.d random matrices via random rounding. arXiv: 1811.07038.

[12] G. Livshyts, K. Tikhomirov, and R. Vershynin. The smallest singular value of inhomogeneous square random matrices. arXiv: 1909.04219.

[13] E. Rebrova and K. Tikhomirov. Covering of random ellipsoids, and invertibility of random matrices with i.i.d heavy-tailed entries. *Israel J. Math.* to appear.

[14] M. Rudelson. Invertibility of random matrices: norm of the inverse. *Ann. of Math.*, 168:575–600, 2008.

[15] M. Rudelson and R. Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. of Math.*, 218(2):600–633, 2008.

[16] T. Tao and V. Vu. On random ±1 matrices: singularity and determinant. *Random Structures Algorithms*, 28(1):1–23, 2006.

[17] T. Tao and V. Vu. On the singularity probability of random Bernoulli matrices. *J. Amer. Math. Soc.*, 20(3):603–628, 2007.

[18] T. Tao and V. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. *Ann. of Math.*, 169:595–632, 2009.

[19] K. Tikhomirov. Singularity of random Bernoulli matrices. *Annals of Math.* to appear.

[20] K. Tikhomirov and A. Litvak. Singularity of sparse Bernoulli matrices. arXiv:2004.03131.