The Strong Restricted Isometry Property of Sub-Gaussian Matrices and the Erasure Robustness Property of Gaussian Random Frames

by

Ran Lu

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Applied Mathematics

Department of Mathematical and Statistical Sciences

University of Alberta

Abstract

In this thesis we will study the robustness property of sub-gaussian random matrices. We first show that the nearly isometry property will still hold with high probability if we erase a certain portion of rows from a sub-gaussian matrix, and we will estimate the erasure ratio with a given small distortion rate in the norm. With this, we establish the strong restricted isometry property (SRIP) and the robust version of Johnson-Lindenstrauss (JL) Lemma for sub-gaussian matrices, which are essential in compressed sensing with corruptions. Then we fix the erasure ratio and deduce the lower and upper bounds of the norm after a erased sub-gaussian matrix acting on a vector, and in this case we can also obtain the corresponding SRIP and the robust version of JL Lemma. Finally, we study the robustness property of Gaussian random finite frames, we will improve existing results.

Acknowledgements

First I would like to express my sincere gratitude to my supervisor Professor Bin Han and co-supervisor Professor Yau Shu Wong, for their patient help and guidance. Their advice and encouragement always inspire me to develop rigorous and creative thinking, which is extremely important for researching, and helped me refreshing my understandings and tackling difficult tasks.

Moreover, many thanks to professors who taught me courses, and to the supervisory committee members, for what I've learned these two years and their valuable support. Also, thanks to the Department of Mathematical and Statistical Sciences which provides me the chance and support for the master program study.

Last but not least, special thanks to all my family members, friends, and colleagues, for their support and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Consider the following problem: In sensor network, we transmit a signal through independent channels and the centre hub receives observations from each channel for further analysis. In practice, it is typical that some channels fail to send the correct measurements, therefore we can only obtain the corrupted data. To deal with problems like this, methods to reconstruct a signal from sampling observations are of great interest to researchers in engineering fields. In general it is impossible to recover the signal if there is nothing known about the signal or the measurement. However, with prior knowledge to the signal, it is possible to recover the signal with negligible or even zero error. Over the past years, researchers have been refreshing their understandings on the relevancy and practicability of the assumptions, and have developed several powerful tools to study the question.

1.2 Compressed Sensing

In signal processing, *compressed sensing* (CS) is a technique for signal recovery via solving certain linear systems. In compressed sensing *without corruption*, the general acquisition setting is represented as $y = Ax$, where $x \in \mathbb{R}^n$ is the *signal*, $y \in \mathbb{R}^m$ is the *measurement* and $A \in \mathbb{R}^{m\times n}$ is the *sensing matrix* where $\mathbb{R}^{m\times n}$ denotes the set of all $m \times n$ real matrices. In general it is extremely difficult to recover the signal if we know nothing about x and A. Around 2004, it was shown by Candés, Romberg and Tao in [6] and Donoho in [10] that if a signal satisfies certain sparsity conditions plus additional assumptions on the sensing matrix, then the reconstruction can be accomplished. The ideas of the proofs in these papers provided the foundations of CS.

The most important concept introduced in compressed sensing is the *restricted isometry property* (RIP), which gives a characterization of the almost norm preserving property of a matrix. This property was first set up by Candés and Tao in [7], and it is defined as follows:

Definition 1.2.1 (Restricted isometry property). We say that a matrix $A \in \mathbb{R}^{m\times n}$ satisfies the *restricted isometry property* (RIP) of order k if there exists $\delta \in [0, 1)$ such that

$$(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2$$

for all $x \in \mathbb{R}^n$ with $\|x\|_0 \leq k$ where $\|x\|_0$ denotes the number of non-zero components of x. The *RIP constant* is defined as

$$\delta_k = \inf\{\delta \in [0,1) : (1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2, \quad \forall x \in \mathbb{R}^n, \|x\|_0 \leq k\}.$$

It turns out that one can exactly recover the signal if a signal is sufficiently sparse and the sensing matrix A satisfies certain restricted isometry conditions. The main theorem was established by Candés and Tao in [7]:

Theorem 1.2.2. *([7, Theorem 1.4]) Let $A \in \mathbb{R}^{m\times n}$ and $k \leq \frac{n}{3}$ be such that the RIP constants satisfy*

$$\delta_k + \delta_{2k} + \delta_{3k} < 1.$$

Let $x_0 \in \mathbb{R}^n$ with $\|x_0\|_0 \leq k$ and put $y = Ax_0$. Then one can exactly recover x_0 via solving the following ℓ^1-minimization problem:

$$(P1) \quad \min_{x\in\mathbb{R}^n} \|x\|_1 \text{ subjected to } Ax = y.$$

In practice, researchers are interested to know which matrices satisfy RIP with small RIP constants. Here are some typical examples:

1. If $A \in \mathbb{R}^{m\times n}(m < n)$ is a random matrix with i.i.d. Gaussian entries with mean zero and variance $\frac{1}{m}$, then the condition for Theorem 1.2.2 holds with overwhelming probability if $k \leq O(m/(\ln(n/m) + 1))$. See e.g. [6] for details.

2. If $A \in \mathbb{R}^{m\times n}(m < n)$ is a matrix with rows randomly chosen from a discrete Fourier transform matrix, then the condition for Theorem 1.2.2 holds with overwhelming probability if $k \leq O(m/\ln n)$. See e.g. [5] for details.

Later in [2], the author proved the generalized result which connects the RIP with the concentration of measure phenomenon. Given a random matrix $A \in \mathbb{R}^{m\times n}$, we say that A is *strongly concentrated around its expectation* if

$$\mathbb{P}\{|\|Ax\|_2 - \|x\|_2| > \epsilon\|x\|_2\} < 2e^{-c_0(\epsilon)m}, \quad \forall x \in \mathbb{R}^n, \epsilon \in (0,1), \tag{1.1}$$

where $c_0(\epsilon) > 0$ depends only on ϵ. One important class of random matrices obeying this property is those with i.i.d. sub-gaussian entries(the formal definition of a sub-gaussian random variable will be given in the next chapter). The general result is as follows:

Theorem 1.2.3. *([2, Theorem 5.1]) Let $A \in \mathbb{R}^{m\times n}$ be a random matrix which is strongly concentrated around its expectation. Then for every $T \subseteq \{1,\ldots,n\}$ with $|T| < n$ and $\delta \in (0,1)$:*

$$(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2, \quad \forall x \in \mathbb{R}^n \text{ with } \operatorname{supp}(x) \subseteq T$$

holds with probability at least

$$1 - 2(12/\delta)^{|T|}e^{-c_0(\delta/2)m},$$

where $|T|$ denotes the cardinality of T and $\operatorname{supp}(x) = \{i \in \{1,\ldots,n\} : x_i \neq 0\}$.

Thus one can conclude that sub-gaussian matrices are good restricted isometries with high probability. It is also interested to consider the RIP for non sub-gaussian matrices. There are several literatures regarding matrices with non sub-gaussian rows or columns, and further details related to the RIP of these matrices are discussed, see e.g. [1], [13], [20], [21].

The RIP is also closely related to the well-known *Johnson-Lindenstrauss(JL) Lemma.* The JL Lemma is about embedding a discrete set of points in a high dimensional Euclidean space into a low dimensional space in a way such that the distances between points are nearly preserved. The formal statement of the lemma is given as follows ([2, Lemma 4.1]):

Lemma 1.2.4 (**Johnson-Lindenstrauss Lemma**). *Let $\epsilon \in (0,1)$. For every finite set Q of points in $\mathbb{R}^N$, let n be a positive integer with $n > O(\ln|Q|/\epsilon^2)$, there* *exists a Lipschitz function* $f : \mathbb{R}^N \to \mathbb{R}^n$ *such that*

$$(1-\epsilon)\|u-v\|^2_{\ell^2(\mathbb{R}^N)} \le \|f(u)-f(v)\|^2_{\ell^2(\mathbb{R}^n)} \le (1+\epsilon)\|u-v\|^2_{\ell^2(\mathbb{R}^N)}$$

for all $u, v \in Q$.

In [1], it was shown that with high probability, f is a linear map with matrix representation $\Phi \in \mathbb{R}^{n\times N}$ such that all entries of Φ are drawn from a random variable with certain moment conditions, see [1] for details of constructing such a matrix. A sketch of the proof of the JL Lemma is the follows: we first prove that the random matrix Φ satisfies two conditions:

- Φ is isotropic, i.e. $\mathbb{E}(\|\Phi x\|^2_{\ell^2(\mathbb{R}^n)}) = \|x\|^2_{\ell^2(\mathbb{R}^N)}$, $\forall x \in \mathbb{R}^N$.
- Φ is strongly concentrated around its expectation.

Then applying a union bound argument over all possible pairs of points in Q yields the result. It is basically the similarity of the RIP and the JL Lemma which suggests the connection. The following theorem provides explicit relations between the JL Lemma and the RIP:

Theorem 1.2.5. *([2, Theorem 5.2]) Suppose that* $n, N \in \mathbb{N}$ *and* $\delta \in (0,1)$ *are given. If* $\Phi \in \mathbb{R}^{n\times N}$ *is a random matrix which is strongly concentrated around its expectation, then there exist* $c_1, c_2 > 0$ *depending only on* δ *such that the RIP holds for* Φ *with the prescribed* δ *and is of order* k *obeying* $k \le c_1 n/\ln(N/k)$ *with probability at least* $1-2e^{-c_2 n}$.

Thus from the theorem we see that JL Lemma implies the RIP.

1.3 Robustness Properties of Random Matrices

A natural generalization of CS is *CS with corruptions*. In this case some elements of the measurement are corrupted. We formulate the model in this setting as follows: Given a signal $x \in \mathbb{R}^n$ with certain sparsity condition and a random projection matrix $A \in \mathbb{R}^{m \times n}$. We only receive $A_T x$ as the observation where $T \subseteq \{1, \ldots, m\}$ is unknown and A_T denotes the sub-matrix by keeping the rows of A with indices in T. In order to reconstruct the signal accurately and efficiently in this case, it is important to verify whether A_T satisfies the RIP. This has led to the concept of the *strong restricted isometry property* (SRIP), which characterizes the norm preserving property of matrices under certain erasure of rows and plays a central role in the study of robustness of matrices. The definition of the SRIP is given in [23]:

Definition 1.3.1. A matrix $A \in \mathbb{R}^{m \times n}$ is said to satisfy the *strong restricted isometry property* (SRIP) of order s and level $[\theta, \omega, \beta]$ with $0 < \theta \leq \omega < 2$ and $\beta \in [0, 1)$ if

$$\theta \|x\|_2^2 \leq \|A_T x\|_2^2 \leq \omega \|x\|_2^2$$

holds for all $x \in \mathbb{R}^n$ with $\|x\|_0 \leq s$ and all $T \subseteq \{1, \ldots, m\}$ with $|T^c| \leq \beta m$.

In [23] the robustness and the SRIP for matrices with Gaussian entries have been settled. In [15] further results about the robustness properties of Gaussian random matrices are proved, and they are used as tools for studying the robust versions of the JL Lemma and the RIP. It is still of great interest to study whether any other types of matrices satisfying the SRIP.
Another area closely connected to CS with corruptions is finite frame theory. In data processing, one commonly used technique is to decompose a given signal

$x \in \mathbb{R}^n$ via the following map

$$x \mapsto \{\langle x, f_k \rangle\}_{k=1}^m,$$

where $\{f_k\}_{k=1}^m$ with $m \geq n$ is a frame in $\mathbb{R}^n$, i.e. $\text{span}\{f_k : 1 \leq k \leq m\} = \mathbb{R}^n$. In this interpretation, we can create a more sparse representation of the original signal, and it might be easier to deal with. Now we consider the following problem: when transmitting signals through a channel with noise and adversarial erasures, it is natural to design a frame which is numerically stable so that it is possible to recover the signal even some data were lost or corrupted. To meet the requirements, we need to design a *numerically erasure robust frame* (NERF), which is a finite frame with bounded condition numbers under a fixed erasure ratio. This concept was introduced in [11]: Let $F \in \mathbb{R}^{m \times n} (m > n)$ be a finite frame such that any n rows form a basis of $\mathbb{R}^n$. Its *condition number* is defined as

$$Cond(F) = \frac{s_{\max}(F)}{s_{\min}(F)},$$

where $s_{\max}(F)$ and $s_{\min}(F)$ denote the largest and smallest singular values of F respectively. For $\beta \in (0, 1)$ such that $\beta m \in \mathbb{N}$, if there exists $C > 0$ such that

$$\max_{T \subseteq \{1,\dots,m\}, |T^c| = \beta m} Cond(F_T) \leq C,$$

then F is called a *numerically erasure robust frame* (NERF) of level (β, C).
It turns out that the study of finite frames shares many similarities with CS with corruptions. Now we are interested in examples of NERF. In [11], the author proved that if a finite frame has entries drawn from i.i.d. Gaussian distributions and its size satisfies certain condition, then this frame is a NERF with

overwhelming probability. Later in [24], it was claimed that the matrix size condition can be omitted thus the result about Gaussian random frames from [11] can be improved. Although the result claimed in [24] is true, but there are problems in the author's argument. The author tried to prove the following result which plays an essential role in the estimation of the smallest singular value of the Gaussian matrix:

Theorem 1.3.2. *([24, Theorem 1.2]) Let $A \in \mathbb{R}^{N\times n} (N > n)$ whose entries are drawn independently from the standard normal distribution and $\eta = \frac{N}{n} > 1$. Then for any $\mu > 0$ there exists $c > 0$ such that*

$$\mathbb{P}(s_{\min}(A) < c\sqrt{n}) < 2e^{-\mu n}.$$

To prove Theorem 1.3.2, the author applied an $\epsilon-$net argument and obtained

$$\mathbb{P}(s_{\min}(A) < c\sqrt{n}) \leq (1+2\epsilon^{-1})^n \left(\frac{2e(c+\epsilon C)}{\eta}\right)^{\frac{N}{2}} + (1+2\epsilon^{-1})^n e^{-\frac{(C-1-\sqrt{\eta})^2}{2}n}, \quad (1.2)$$

where $C > 0$ depends on η and $\epsilon \in (0,1)$. In order to control the right hand side of (1.2) by $2e^{-\mu n}$, the author claimed that it suffices to have

$$-\mu \geq \ln(1+2\epsilon^{-1}) + \frac{\eta}{2}[\ln(2e) - \ln\eta + 2\ln(c+\epsilon C)], \quad (1.3)$$

$$-\mu \geq -\frac{1}{2}(C-1-\sqrt{\eta})^2, \quad (1.4)$$

which is problematic because in fact from (1.3) and (1.4) one cannot conclude that the right-hand side of (1.2) is bounded by $2e^{-\mu n}$. However if we replace the right-hand side of (1.4) by

$$-\mu \geq \ln(1+2\epsilon^{-1}) - \frac{1}{2}(C-1-\sqrt{\eta})^2, \tag{1.5}$$

then one can conclude that $2e^{-\mu n}$ is an upper bound of the right-hand side of (1.2) from (1.3) and (1.5). Therefore, although using the author's approach one can prove Theorem 1.3.2, but the result will be significantly weaker, i.e. the number c in Theorem 1.3.2 will be much smaller than the one which was provided by the author, and hence the bounds of condition numbers will be much worse.

1.4 Main Work and Thesis Structure

In this thesis, we discuss the robustness property of sub-gaussian random matrices. Sub-gaussian matrices share several crucial properties which we have mentioned in previous sections:

- If a matrix has entries drawn from i.i.d. sub-gaussian random variables, then this matrix satisfies the RIP with high probability.
- A sub-gaussian random matrix satisfies the concentration inequality (1.1).

Thus sub-gaussian matrices appear frequently in the study of CS. The most special case is that a matrix has entries drawn from i.i.d. Gaussian distributions. In the past years, Gaussian random matrices have been studied thoroughly. Therefore, researchers are interested to know whether it is possible to obtain similar results for sub-gaussian cases.

In this thesis, we will extend certain results from the Gaussian case to sub-gaussian cases. The structure of the thesis is organized as follows: Chapter 2 recalls the basic concepts and properties of sub-gaussian random variables, including concentration inequalities and order statistics results which are useful in our study. Chapter 3 will provide the study of robustness properties of sub-gaussian random matrices. We will focus on the following question: how large can the erasure ratio $\beta \in (0, 1)$ be so that a normalized Gaussian random matrix $A \in \mathbb{R}^{m\times n}$ with βm arbitrarily erased rows has the nearly isometry property with high probability. In Chapter 4 we will establish the SRIP and the robust JL Lemma for sub-gaussian matrices. In particular we will see that the results can be further improved in the case of Bernoulli matrices. Chapter 5 contains the results related to Gaussian finite frames, we will improve the results in [24] and provide the implicit forms of the bounds on condition numbers of the reduced matrices. Finally a summary and further discussion are provided in Chapter 6.

Chapter 2

Sub-gaussian Random Variables

In this chapter we provide some preliminary background related to sub-gaussian random variables, including characterizations and the concentration of measure properties, which play important roles in our study of the robustness of sub-gaussian random matrices and finite frames.

2.1 Basics

All definitions and results provided in this section are known (see e.g. [4]).

Definition 2.1.1. A real random variable X is called $b-$*sub-gaussian* for some $b > 0$ if

$$\mathbb{E}(e^{tX}) \leq e^{\frac{b^2t^2}{2}}, \quad \forall t \in \mathbb{R}. \tag{2.1}$$

Notation: $X \sim \mathrm{Sub}(b^2)$. We will always assume that b is the smallest positive number so that (2.1) holds, and in this case b^2 is called the *sub-gaussian* moment of X.

We have the following examples of sub-gaussian random variables:

Example 2.1.2. (i) **Gaussian:** If X is a standard normal random variable, then X is $1-$sub-gaussian.

(ii) **Symmetric Bernoulli:** If X satisfies

$$\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = \frac{1}{2},$$

then X is $1-$sub-gaussian.

(iii) **Bounded centred radom variables :** If X is a random variable such that there exists $b > 0$ with $|X| \leq b$ and $\mathbb{E}(X) = 0$, then $X \sim \mathrm{Sub}(b^2)$.

Next we provide the following basic properties of sub-gaussian random variables:

Proposition 2.1.3. *(i) If $X \sim \mathrm{Sub}(b^2)$, then X is centred, i.e. $\mathbb{E}(X) = 0$.*

(ii) If $X \sim \mathrm{Sub}(b^2)$, then $\sigma_X^2 = \mathrm{Var}(X) \leq b^2$.

(iii) If $X \sim \mathrm{Sub}(b^2)$, then for all $c \in \mathbb{R}$ we have $cX \sim \mathrm{Sub}(c^2 b^2)$.

(iv) If $X_1 \sim \mathrm{Sub}(b_1^2)$ and $X_2 \sim \mathrm{Sub}(b_2^2)$, then $X_1 + X_2 \sim \mathrm{Sub}((b_1 + b_2)^2)$. Moreover, if X_1, X_2 are independent, then $X_1 + X_2 \sim \mathrm{Sub}(b_1^2 + b_2^2)$.

(v) If $X = (X_1, X_2, \ldots, X_N)$ is a random vector such that X_i's are i.i.d. random variables with $X_i \sim \mathrm{Sub}(b^2)$, then for any $\alpha \in \mathbb{R}^N$ we have

$$\langle X, \alpha \rangle \sim \mathrm{Sub}(\|\alpha\|_2^2 b^2).$$

Characterization of sub-gaussian random variables: We can define sub-gaussian random variables in the following equivalent ways:

Theorem 2.1.4. *The following are equivalent for a centred random variable X:*

(i) $X \sim \mathrm{Sub}(b^2)$ *for some* $b > 0$*.*

(ii) *For all* $\lambda > 0$*, there exists* $c > 0$ *such that* $\mathbb{P}(|X| > \lambda) \leq 2e^{-c\lambda^2}$*. In fact we may choose* $c = \frac{1}{2b^2}$*.*

(iii) *For every* $\xi > 1$*, there exists* $a > 0$ *depending on* ξ *such that* $\mathbb{E}(e^{aX^2}) \leq \xi$*. In fact, we can choose* $a = \frac{\xi - 1}{2b^2(\xi+1)}$*.*

2.2 Concentration of Measure Phenomena

The concentration of measure phenomena is extremely important in the study of measure and probability theory, and has various applications and consequences in areas such as Banach space theory and random matrix theory. The concentration of measure for Gaussian matrices has been studied thoroughly in the past (see e.g. [17]). The most useful result is the Lipschitz concentration for Gaussian matrices, which has several applications in the study of CS. In this section we will develop the Lipschitz concentration property for sub-gaussian matrices.

First we recall the concentration of measure for sub-gaussian random matrices:

Theorem 2.2.1 (Concentration of measure for sub-gaussian matrices). *([9, Lemma 6.1]) Suppose that* $A \in \mathbb{R}^{m \times n}$ *is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* $\frac{1}{m}$*. Then for all* $x \in \mathbb{R}^n$*, we have:*

(i) $\mathbb{E}(\|Ax\|_2^2) = \|x\|_2^2$*.*

(ii) *There exists* $\kappa > 0$ *(depending on the distribution) such that for any* $\epsilon \in (0, 1)$ *we have*

$$\mathbb{P}(\|Ax\|_2^2 - \|x\|_2^2 > \epsilon \|x\|_2^2) < \exp(-\kappa \epsilon^2 m),$$

$$\mathbb{P}(\|Ax\|_2^2 - \|x\|_2^2 < -\epsilon\|x\|_2^2) < \exp(-\kappa\epsilon^2 m).$$

Remark 2.2.2. From [1], if $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. standard normal or symmetric Bernoulli entries, then for all $x \in \mathbb{R}^n$ we have

$$\mathbb{P}\left(\frac{1}{m}\|Ax\|_2^2 - \|x\|_2^2 > \epsilon\|x\|_2^2\right) < \exp\left(-\left(\frac{\epsilon^2}{4} - \frac{\epsilon^3}{6}\right)m\right),$$

$$\mathbb{P}\left(\frac{1}{m}\|Ax\|_2^2 - \|x\|_2^2 < -\epsilon\|x\|_2^2\right) < \exp\left(-\left(\frac{\epsilon^2}{4} - \frac{\epsilon^3}{6}\right)m\right).$$

Thus in this case we may take $\kappa = \frac{1}{12}$.

We now prove the Lipschitz concentration inequality for sub-gaussian random variables. We borrow the elegant arguments used in the proof of [16, Proposition 2.3]:

Theorem 2.2.3 (Lipschitz concentration inequality for sub-gaussian distributions). *Let $X = (X_1, \ldots, X_m)$ has i.i.d. entries with $X_i \sim \mathrm{Sub}(b^2)$. Let $f : \mathbb{R}^m \to \mathbb{R}$ be a 1−Lipschitz function, i.e. $|f(x) - f(y)| \leq \|x - y\|_2$, $\forall x, y \in \mathbb{R}^m$. Then for all $t > 0$ we have*

$$\mathbb{P}(f(X) - \mathbb{E}(f(X)) \geq t) \leq \exp\left(-\frac{t^2}{5b^2}\right),$$

$$\mathbb{P}(f(X) - \mathbb{E}(f(X)) \leq -t) \leq \exp\left(-\frac{t^2}{5b^2}\right).$$

Proof. By Rademacher's theorem a Lipschitz map is differentiable almost everywhere, so it suffices to prove the result for differentiable f. As f has Lipschitz constant 1, we have $\|\nabla f\|_2 \leq 1$. With out loss of generality assume $\mathbb{E}(f(X)) = 0$.

Let X' be an independent copy of X, and let $\gamma : [0, 1] \to \mathbb{R}^m$ be a smooth path connecting X and X' with

$$\gamma(t) = X\cos\left(\frac{\pi}{2}t\right) + X'\sin\left(\frac{\pi}{2}t\right).$$

Then

$$\gamma'(t) = \frac{\pi}{2}\left(X\sin\left(\frac{\pi}{2}t\right) + X'\cos\left(\frac{\pi}{2}t\right)\right) =: \frac{\pi}{2}Y(t).$$

Observe that Y has i.i.d. components with $Y_i \sim \mathrm{Sub}(b^2)$. By fundamental theorem of line integral we have

$$f(X) - f(X') = \frac{\pi}{2}\int_0^1 \langle \nabla f(\gamma(t)), Y(t)\rangle \,\mathrm{d}t.$$

Therefore by Jensen's inequality and Fubini's theorem:

$$\mathbb{E}(\exp(\lambda(f(X) - f(X'))) \leq \int_0^1 \mathbb{E}\exp(\frac{\pi}{2}\lambda\langle \nabla f(\gamma(t)), Y(t)\rangle))\,\mathrm{d}t.$$

As $\|\nabla f\|_2 \leq 1$, and Y has i.i.d. $\mathrm{Sub}(b^2)$ entries, it follows that

$$\langle \nabla f(\gamma(t)), Y(t)\rangle \sim \mathrm{Sub}(b^2).$$

Thus

$$\mathbb{E}(\exp(\frac{\pi}{2}\lambda\langle \nabla f(\gamma(t)), Y(t)\rangle)) \leq \exp\left(\frac{b^2\lambda^2\pi^2}{8}\right).$$

Therefore

$$\mathbb{E}(\exp(\lambda(f(X) - f(X'))) \leq \exp\left(\frac{b^2\lambda^2\pi^2}{8}\right).$$

Note that as X and X' are i.i.d. copies, so we have $\mathbb{E}(f(X)) = \mathbb{E}(f(X')) = 0$, and

Jensen's inequality yields

$$\mathbb{E}(\exp(\lambda f(X'))) \geq \exp(\mathbb{E}(\lambda f(X'))) = 1, \quad \forall \lambda \in \mathbb{R}.$$

Therefore

$$\mathbb{E}(\exp(\lambda f(X))) \leq \mathbb{E}(\exp(\lambda(f(X) - f(X'))) \leq \exp\left(\frac{b^2\lambda^2\pi^2}{8}\right), \quad \forall \lambda > 0.$$

Thus for all $\lambda, t > 0$, we have

$$\mathbb{P}(f(X) \geq t) = \mathbb{P}(\exp(\lambda f(X)) \geq e^{\lambda t}) \leq \frac{\mathbb{E}(\exp(\lambda f(X)))}{e^{\lambda t}} \leq \exp\left(\frac{b^2\lambda^2\pi^2}{8} - \lambda t\right).$$

By setting $\lambda = \frac{4t}{b^2\pi^2}\left(1 + \sqrt{1 - \frac{\pi^2}{10}}\right)$, we have

$$\mathbb{P}(f(x) \geq t) \leq \exp\left(-\frac{t^2}{5b^2}\right).$$

Hence the proof is complete. □

2.3 Order Statistics

Suppose $y_1, \ldots, y_m$ are i.i.d. random variables such that $y_j \sim \mathrm{Sub}(b^2)$. Let $S \subseteq \{1, \ldots, m\}$. Define

$$F_S : \mathbb{R}^m \to \mathbb{R}, \quad x \mapsto \sqrt{\sum_{j \in S} x_{(j)}^2},$$

where $x_{(1)}, \ldots, x_{(|S|)}$ denotes the non-increasing rearrangements of elements in S in magnitudes, i.e. $|x_{(1)}| \geq \cdots \geq |x_{(|S|)}|$. Then it is easy to see that F_S is

1−Lipschitz. So for all $t > 0$, it follows that

$$
\begin{aligned}
&\mathbb{P}\left(\sqrt{\frac{1}{|S|}\sum_{j\in S} y_{(j)}^2} \geq t + \mathbb{E}\sqrt{\frac{1}{|S|}\sum_{j\in S} y_{(j)}^2}\right) \\
=&\mathbb{P}\left(\sqrt{\sum_{j\in S} y_{(j)}^2} \geq t\sqrt{|S|} + \mathbb{E}\sqrt{\sum_{j\in S} y_{(j)}^2}\right) \\
\leq &\exp\left(-\frac{t^2|S|}{5b^2}\right).
\end{aligned}
$$

Corollary 2.3.1. *Suppose that $y_1, \dots, y_m$ are i.i.d. random variables with $y_i \sim \mathrm{Sub}(b^2)$. Let $y_{(1)}, \dots, y_{(m)}$ be non-increasing rearrangements of y_i's in magnitudes, i.e. $|y_{(1)}| \geq \cdots \geq |y_{(m)}|$. Then*

$$
\mathbb{E}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2}\right) \leq \sqrt{2eb^2 \ln \frac{em}{k}}.
$$

Proof. Let $\xi > 1$ and choose $t = \frac{\xi-1}{2b^2(\xi+1)}$. It follows that:

$$
\begin{aligned}
\exp\left(\mathbb{E}\left(\frac{1}{k}\sum_{j=1}^{k} ty_{(j)}^2\right)\right) &= \exp\left(\frac{1}{k}\sum_{j=1}^{k}\mathbb{E}(ty_{(j)}^2)\right) \leq \frac{1}{k}\sum_{j=1}^{k}\exp(\mathbb{E}(ty_{(j)}^2)) \\
&\leq \frac{1}{k}\sum_{j=1}^{k}\mathbb{E}(\exp(ty_{(j)}^2)) \leq \frac{1}{k}\sum_{j=1}^{m}\mathbb{E}(\exp(ty_{(j)}^2)) \\
&= \frac{1}{k}\mathbb{E}\left(\sum_{j=1}^{m}\exp(ty_{(j)}^2)\right) = \frac{1}{k}\mathbb{E}\left(\sum_{j=1}^{m}\exp(ty_j^2)\right) \\
&= \frac{1}{k}\sum_{j=1}^{m}\mathbb{E}(\exp(ty_j^2)) \leq \frac{1}{k}\sum_{j=1}^{m}\xi \quad \text{(by (iii) of Theorem 2.1.4)} \\
&= \xi\frac{m}{k}.
\end{aligned}
$$

Taking logarithms on both sides and using Jensen's inequality, we have

$$\mathbb{E}\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} \leq \sqrt{\frac{2b^2(\xi+1)}{\xi-1}\ln\frac{\xi m}{k}}.$$

Setting $\xi = e$ yields the result. $\square$

Chapter 3

Sub-gaussian Matrices under Arbitrary Erasure of Rows with Small Distortion in Norms

In this chapter we are going to study the robustness property of sub-gaussian matrices with a small distortion rate in norms. Our main goal is to establish the robust version of the JL Lemma and the SRIP with a small distortion in norms. Throughout this chapter, $y_{(1)}, \ldots, y_{(m)}$ denote the non-increasing rearrangements of the sequence of random variables $y_1, \ldots, y_m$.

The corresponding results in this chapter for the Gaussian case have been settled in [15]. We will extend the results in [15] from Gaussian random matrices to sub-gaussian random matrices.

Main contributions: Extending the robust version of the JL Lemma and the RIP from the Gaussian case to sub-gaussian cases.

3.1 Estimate of the Erasure Ratio

Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\mathrm{Sub}(b^2)$ and has variance 1. Let $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$. Then Ax is a random vector with i.i.d. components such that each component obeys $\mathrm{Sub}(b^2)$. For $\epsilon \in (0,1)$, $\beta \in [0,1)$, define

$$\Omega_{\epsilon,\beta} = \left\{ \left| \frac{1}{|T|}\|A_T x\|_2^2 - 1 \right| \leq \epsilon \text{ for all } T \subseteq \{1,\dots,m\} \text{ such that } |T^c| \leq \beta m \right\}.$$

From the concentration inequality for the sub-gaussian distributions we have

$$\mathbb{P}\left(\left| \frac{1}{m}\|Ax_0\|_2^2 - 1 \right| > \epsilon \right) \leq 2\exp\left(-\kappa\varepsilon^2 m\right).$$

For $\epsilon \in (0,1)$ and $\alpha \in (0,1]$, set

$$\beta_{\epsilon,\alpha} := \sup\{\beta \in [0,1) : \mathbb{P}(\Omega_{\epsilon,\beta}) \geq 1 - 3e^{-\alpha\kappa\epsilon^2 m}, \quad \forall m \in \mathbb{N}\}.$$

By studying the quantity $\beta_{\epsilon,\alpha}$, we will know how large the erasure ratio can be under a certain small distortion rate in norms.

Remark 3.1.1. For $\gamma \in [0,1)$ with $\gamma m \in \mathbb{N}$, define

$$T_\gamma := \{T \subseteq \{1,\dots,m\} : |T^c| = \gamma m\}.$$

Then it is easy to see that

$$\min_{T\in T_\beta} \frac{1}{|T|}\|Ax\|_2^2 \leq \min_{T\in T_\gamma} \frac{1}{|T|}\|Ax\|_2^2 \leq \max_{T\in T_\gamma} \frac{1}{|T|}\|Ax\|_2^2 \leq \max_{T\in T_\beta} \frac{1}{|T|}\|Ax\|_2^2,$$

for $0 \leq \gamma \leq \beta < 1$ with $\gamma m, \beta m \in \mathbb{N}$.

Let $y = Ax$. For $\beta \in [0,1)$, define $\gamma := \lfloor \beta m \rfloor / m$. Then $\gamma m = k$ is an integer and $0 \leq \gamma \leq \beta$. Let y_T denotes the sub-vector by keeping components of y with indices in T. It follows that

$$\Omega_{\epsilon,\beta} = \left\{ 1 - \epsilon \leq \min_{T \in T_\gamma} \frac{\|y_T\|_2^2}{|T|} \leq \max_{T \in T_\gamma} \frac{\|y_T\|_2^2}{|T|} \leq 1 + \epsilon \right\}.$$

Now we estimate the maximum erasure ration $\beta_{\epsilon,\alpha}$.

Lemma 3.1.2. *Suppose that $A \in \mathbb{R}^{m \times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\mathrm{Sub}(b^2)$ and has variance 1. Let $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$ and let $y = Ax$. For $0 < \alpha < 1$ and $0 < \epsilon \leq \min\left(1, \frac{1-\sqrt{\alpha}}{5b^2\alpha\kappa}\right)$, if*

$$0 < \beta \leq \frac{1-\sqrt{\alpha}}{1+\epsilon}\epsilon \qquad \textit{and} \qquad 0 < \beta \ln \beta \leq \frac{\epsilon}{2eb^2}\left(\sqrt{1-\sqrt{\alpha}} - \sqrt{5b^2\alpha\kappa\epsilon}\right)^2, \quad (3.1)$$

then

$$\mathbb{P}(\Omega_{\epsilon,\beta}) \geq 1 - 3\exp\left(-\alpha\kappa\epsilon^2 m\right).$$

Proof. We borrow the argument from the proof of [15, Lemma 3.1]. It follows from Remark 3.1.1 that

$$\begin{aligned}
\mathbb{P}(\Omega_{\epsilon,\beta}) &= \mathbb{P}\left\{ 1 - \epsilon \leq \frac{\|y\|_2^2 - (y_{(1)}^2 + \cdots + y_{(k)}^2)}{m-k} \leq \frac{\|y\|_2^2 - (y_{(m-k+1)}^2 + \cdots + y_{(m)}^2)}{m-k} \leq 1 + \epsilon \right\} \\
&= \mathbb{P}(\left\{ \frac{y_{(1)}^2 + \cdots + y_{(k)}^2}{k} \leq \frac{\|y\|_2^2}{k} - \frac{(1-\gamma)(1-\epsilon)}{\gamma} \right\} \\
&\qquad \cap \left\{ \frac{y_{(m-k+1)}^2 + \cdots + y_{(m)}^2}{k} \geq \frac{\|y\|_2^2}{k} - \frac{(1-\gamma)(1+\epsilon)}{\gamma} \right\}) \\
&\geq 1 - (\mathbb{P}(B_0^c) + \mathbb{P}(B_1^c) + \mathbb{P}(B_2^c)),
\end{aligned}$$

where

$$B_0 := \left\{1 - \sqrt{\alpha\epsilon} \leq \frac{\|y\|_2^2}{m} \leq 1 + \sqrt{\alpha\epsilon}\right\},$$

$$B_1 := \left\{\frac{y_{(1)}^2 + \cdots + y_{(k)}^2}{k} \leq 1 - \epsilon + \frac{1-\sqrt{\alpha}}{\gamma}\epsilon\right\},$$

$$B_2 := \left\{\frac{y_{(m-k+1)}^2 + \cdots + y_{(m)}^2}{k} \geq 1 + \epsilon - \frac{1-\sqrt{\alpha}}{\gamma}\epsilon\right\}.$$

Now we need

$$\mathbb{P}(B_0^c) + \mathbb{P}(B_1^c) + \mathbb{P}(B_2^c) \leq 3\exp\left(-\alpha\kappa\epsilon^2 m\right).$$

We will estimate three probabilities separately:

- **Estimating** $\mathbb{P}(B_1^c)$: For all $t > 0$, by Corollary 2.3.1 and Theorem 2.2.3, we have:

$$\mathbb{P}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} - \sqrt{2eb^2\ln\frac{e}{\gamma}} > t\right) \leq \mathbb{P}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} - \mathbb{E}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2}\right) > t\right)$$
$$\leq \exp\left(-\frac{t^2\gamma m}{5b^2}\right).$$

Set

$$t := \sqrt{1 - \epsilon + \frac{1-\sqrt{\alpha}}{\gamma}\epsilon} - \sqrt{2eb^2\ln\frac{e}{\gamma}}.$$

Observe that by the assumption $\epsilon \leq \frac{1-\sqrt{\alpha}}{5b^2\alpha\kappa}$ we have

$$f_{\epsilon,\alpha} := \sqrt{\frac{\epsilon}{2}}\left(\sqrt{1-\sqrt{\alpha}} - \sqrt{5b^2\alpha\kappa\epsilon}\right) \geq 0.$$

Now consider the function

$$g(x) = 2x^2 + 2\sqrt{5b^2\alpha\kappa\epsilon^2}x.$$

It is easy to see that

$$g(x) \le g(f_{\epsilon,\alpha}) = (1-\sqrt{\alpha})\epsilon - 5b^2\alpha\kappa\epsilon^2, \quad \forall x \in [0, f_{\epsilon,\alpha}]. \tag{3.2}$$

On the other hand the function $h(x) = x \ln \frac{e}{x}$ is increasing on $(0, 1]$. Therefore

$$0 < \sqrt{eb^2\gamma \ln \frac{e}{\gamma}} \le \sqrt{eb^2\beta \ln \frac{e}{\beta}} \le f_{\epsilon,\alpha}.$$

Plugging $x = \sqrt{eb^2\gamma \ln \frac{e}{\gamma}}$ into (3.2) we get

$$2eb^2\gamma \ln \frac{e}{\gamma} + 2\sqrt{5b^2\alpha\kappa\epsilon^2}\sqrt{eb^2\gamma \ln \frac{e}{\gamma}} \le (1-\sqrt{\alpha})\epsilon - 5b^2\alpha\kappa\epsilon^2. \tag{3.3}$$

Dividing both sides of (3.3) by γ and rearranging terms gives

$$\begin{aligned}\frac{1-\sqrt{\alpha}}{\gamma}\epsilon &\ge 2eb^2 \ln \frac{e}{\gamma} + 2\sqrt{\frac{5b^2}{\gamma}\alpha\kappa\epsilon^2}\sqrt{2eb^2 \ln \frac{e}{\gamma}} + \frac{5b^2}{\gamma}\alpha\kappa\epsilon^2 \\ &= \left(\sqrt{2eb^2 \ln \frac{e}{\gamma}} + \sqrt{\frac{5b^2}{\gamma}\alpha\kappa\epsilon^2}\right)^2.\end{aligned}$$

Therefore

$$\sqrt{2eb^2 \ln \frac{e}{\gamma}} + \sqrt{\frac{5b^2}{\gamma}\alpha\kappa\epsilon^2} \le \sqrt{\frac{1-\sqrt{\alpha}}{\gamma}\epsilon} \le \sqrt{1-\epsilon+\frac{1-\sqrt{\alpha}}{\gamma}\epsilon}.$$

Thus

$$t \ge \sqrt{\frac{5b^2}{\gamma}\alpha\kappa\epsilon^2} > 0.$$

It follows that

$$
\begin{aligned}
\mathbb{P}(B_1^c) &= \mathbb{P}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} > \sqrt{1-\epsilon+\frac{1-\sqrt{\alpha}}{\gamma}\epsilon}\right) \\
&\leq \exp\left(-\frac{t^2\gamma m}{5b^2}\right) \\
&\leq \exp\left(-\alpha\kappa\epsilon^2 m\right).
\end{aligned}
$$

- **Estimating** $\mathbb{P}(B_2^c)$: It is easy to see that

$$
1+\epsilon-\frac{1-\sqrt{\alpha}}{\gamma}\epsilon \leq 1+\epsilon-\frac{1-\sqrt{\alpha}}{\beta}\epsilon \leq 1+\epsilon-(1+\epsilon)=0.
$$

  Thus
$$
\mathbb{P}(B_2) = \mathbb{P}\left(\frac{y_{(m-k+1)}^2+\cdots+y_{(m)}^2}{k} \geq 0\right) = 1.
$$

  Hence
$$
\mathbb{P}(B_2^c) = 0.
$$

- **Estimating** $\mathbb{P}(B_0^c)$: It is trivial from the concentration inequality for sub-gaussian matrices that

$$
\mathbb{P}(B_0^c) < 2\exp\left(-\alpha\kappa\epsilon^2 m\right).
$$

Hence combining everything together we have

$$
\mathbb{P}(\Omega_{\epsilon,\beta}) \geq 1-3\exp\left(-\alpha\kappa\epsilon^2 m\right).
$$

Thus the proof is complete. $\square$

Theorem 3.1.3. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *Let $x_0 \in \mathbb{R}^n$ with $\|x_0\|_2 = 1$ and put $y = Ax_0$. For $0 < \alpha < 1$ and $0 < \epsilon \leq \min\left(1, \frac{1-\sqrt{\alpha}}{40b^2\alpha\kappa}\right)$, if*

$$0 < \beta \leq \frac{(1-\sqrt{\alpha})\epsilon}{16\ln\frac{4eb^2}{(1-\sqrt{\alpha})\epsilon}} =: \eta(\epsilon, \alpha),$$

then (3.1) *holds. Moreover:*

$$\beta_{\epsilon,\alpha} \geq \frac{(1-\sqrt{\alpha})\epsilon}{32\ln\frac{1}{\epsilon}}, \quad \forall 0 < \epsilon \leq \frac{(1-\sqrt{\alpha})}{4eb^2}.$$

Proof. We apply the same argument as in the proof of [15, Theorem 3.2]. We have

$$0 < \beta \leq \eta(\epsilon, \alpha) \leq \frac{(1-\sqrt{\alpha})\epsilon}{16\ln(4eb^2)} < \frac{(1-\sqrt{\alpha})\epsilon}{2} \leq \frac{(1-\sqrt{\alpha})\epsilon}{1+\epsilon}.$$

Let $f_{\epsilon,\alpha}$ be the same as in the proof of Lemma 3.1.2, we have

$$f_{\epsilon,\alpha}^2 = \frac{\epsilon}{2}\left(\sqrt{1-\sqrt{\alpha}} - \sqrt{5b^2\alpha\kappa\epsilon}\right)^2 \geq \frac{\epsilon}{2}\left[\left(1-\sqrt{\frac{1}{8}}\right)(\sqrt{1-\sqrt{\alpha}})\right]^2 = \frac{9-4\sqrt{2}}{16}(1-\sqrt{\alpha})\epsilon.$$

For $t \geq \ln 4$ we have

$$\ln(4et) < (8-4\sqrt{2})t. \tag{3.4}$$

By plugging $t = \ln\frac{1}{x}$ with $0 < x \leq \frac{1}{4}$ into (3.4), we get

$$\frac{1}{\ln\frac{1}{x}}\ln\left(4e\ln\frac{1}{x}\right) < 8-4\sqrt{2} \quad\Longrightarrow\quad \frac{x}{4\ln\frac{1}{x}}\ln\frac{4e\ln\frac{1}{x}}{x} < \frac{9-4\sqrt{2}}{4}x.$$

Recall that for a sub-gaussian random variable, its sub-gaussian moment is always great than or equal to its variance. Thus $b^2 \geq 1$. Hence by taking

$$x := \frac{1}{4eb^2}(1-\sqrt{\alpha})\epsilon < \frac{1}{4},$$

we conclude that

$$\begin{aligned}
\eta(\epsilon,\alpha)\ln\frac{e}{\eta(\epsilon,\alpha)} &= \frac{x}{4\ln\frac{1}{x}}\ln\frac{4e\ln\frac{1}{x}}{x} \\
&< \frac{9-4\sqrt{2}}{4}x \\
&= \frac{9-4\sqrt{2}}{16eb^2}(1-\sqrt{\alpha})\epsilon \\
&\leq \frac{f_{\epsilon,\alpha}^2}{eb^2}.
\end{aligned}$$

Therefore as $x\ln\frac{e}{x}$ is increasing on $(0,1]$ we have $\beta\ln\frac{e}{\beta} < \frac{f_{\epsilon,\alpha}^2}{eb^2}$. Hence we've proved that two inequalities in (3.1) hold.

Since $\eta(\epsilon,\alpha) < \frac{1-\sqrt{\alpha}}{1+\epsilon}\epsilon$ and $\eta(\epsilon,\alpha)\ln\frac{e}{\eta(\epsilon,\alpha)} < \frac{f_{\epsilon,\alpha}^2}{eb^2}$, then there exists $\delta > 0$ such that $\eta(\epsilon,\alpha)+\delta$ satisfies (3.1). Thus by definition of $\beta_{\epsilon,\alpha}$:

$$\beta_{\epsilon,\alpha} \geq \eta(\epsilon,\alpha) > \frac{(1-\sqrt{\alpha})\epsilon}{16\left(\ln\frac{4eb^2}{(1-\sqrt{\alpha})\epsilon}+\ln\frac{1}{\epsilon}\right)} \geq \frac{(1-\sqrt{\alpha})\epsilon}{32\ln\frac{1}{\epsilon}}, \quad \forall 0<\epsilon\leq\frac{(1-\sqrt{\alpha})}{4eb^2}.$$

□

3.2 Estimate under the Uniform Normalization

For $\epsilon \in (0,1)$ and $\beta \in [0,1)$, we define the following notion which is closely related to $\Omega_{\epsilon,\beta}$:

$$\tilde{\Omega}_{\epsilon,\beta} = \left\{\left|\frac{1}{m}\|A_T x\|_2^2 - 1\right| \leq \epsilon \text{ for all } T \subseteq \{1,\ldots,m\} \text{ such that } |T^c| \leq \beta m\right\}.$$

In this case, we use the uniform normalization factor $\frac{1}{m}$ instead of $\frac{1}{|T|}$. This means we first normalize the original matrix and then erase rows, whereas in the previous case (where the factor $\frac{1}{|T|}$ is used) we first erase rows and then normalize the reduced matrix.

For $\epsilon \in (0,1)$ and $\alpha > 0$, we may also define the following quantity:

$$\tilde{\beta}_{\epsilon,\alpha} = \sup\{\beta \in [0,1) : \mathbb{P}(\tilde{\Omega}_{\epsilon,\alpha}) \geq 1 - 3\exp(-\alpha\kappa\epsilon^2 m), \quad \forall m \in \mathbb{N}\}.$$

Moreover, we can use the same argument as in the case of $\beta_{\epsilon,\alpha}$ to estimate $\tilde{\beta}_{\epsilon,\alpha}$:

Lemma 3.2.1. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\mathrm{Sub}(b^2)$ and has variance 1. Let $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$ and put $y = Ax$. For $0 < \alpha < 1$ and $0 < \epsilon \leq \min\left(1, \frac{1-\sqrt{\alpha}}{5b^2\alpha\kappa}\right)$, if*

$$0 < \beta \ln \beta \leq \frac{\epsilon}{2eb^2}\left(\sqrt{1-\sqrt{\alpha}} - \sqrt{5b^2\alpha\kappa\epsilon}\right)^2, \tag{3.5}$$

then

$$\mathbb{P}(\tilde{\Omega}_{\epsilon,\beta}) \geq 1 - 3\exp\left(-\alpha\kappa\epsilon^2 m\right).$$

Proof. Set $\gamma = \lfloor \beta m \rfloor / m$ and $k = \gamma m$. It follows that

$$\begin{aligned}
\mathbb{P}(\tilde{\Omega}_{\epsilon,\beta}) &= \mathbb{P}\left\{1 - \epsilon \leq \frac{\|y\|_2^2 - (y_{(1)}^2 + \cdots + y_{(k)}^2)}{m} \leq \frac{\|y\|_2^2}{m} \leq 1 + \epsilon\right\} \\
&\geq \mathbb{P}\left\{y_{(1)}^2 + \cdots + y_{(k)}^2 \geq (1-\sqrt{\alpha})\epsilon m \text{ and } \left|\frac{1}{m}\|y\|_2^2 - 1\right| \leq \sqrt{\alpha}\epsilon\right\} \\
&\geq 1 - \mathbb{P}(B_0^c) - \mathbb{P}(B_3^c),
\end{aligned}$$

where

$$B_0 := \left\{\left|\frac{1}{m}\|y\|_2^2 - 1\right| \leq \sqrt{\alpha}\epsilon\right\},$$

$$B_3 := \left\{ y_{(1)}^2 + \cdots + y_{(k)}^2 \geq (1 - \sqrt{\alpha})\epsilon m \right\}.$$

We want to control $\mathbb{P}(B_0^c) + \mathbb{P}(B_3^c)$ by $3\exp\left(-\alpha\kappa\epsilon^2 m\right)$. Note that

$$\mathbb{E}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2}\right) \leq \sqrt{2eb^2 \ln \frac{e}{\gamma}}.$$

Thus for $t > 0$ we have

$$\begin{aligned} \mathbb{P}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} - \sqrt{2eb^2 \ln \frac{e}{\gamma}} > t\right) &\leq \mathbb{P}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} - \mathbb{E}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2}\right) > t\right) \\ &\leq \exp\left(-\frac{t^2\gamma m}{5b^2}\right). \end{aligned}$$

Set

$$t := \sqrt{\frac{1-\sqrt{\alpha}}{\gamma}\epsilon} - \sqrt{2eb^2 \ln \frac{e}{\gamma}}.$$

Define

$$f_{\epsilon,\alpha} := \sqrt{\frac{\epsilon}{2}}\left(\sqrt{1-\sqrt{\alpha}} - \sqrt{5b^2\alpha\kappa\epsilon}\right) \geq 0.$$

Then from the proof of Lemma 3.1.2, we have

$$0 < \sqrt{eb^2\gamma \ln \frac{e}{\gamma}} \leq \sqrt{eb^2\beta \ln \frac{e}{\beta}} \leq f_{\epsilon,\alpha}.$$

Thus

$$\sqrt{2eb^2 \ln \frac{e}{\gamma}} \leq \sqrt{\frac{1-\sqrt{\alpha}}{\gamma}\epsilon} - \sqrt{\frac{5b^2}{\gamma}\alpha\kappa\epsilon^2} \quad \Longrightarrow \quad t \geq \sqrt{\frac{5b^2}{\gamma}\alpha\kappa\epsilon^2} > 0.$$

It follows that

$$\begin{aligned}\mathbb{P}(B_3^c) &= \mathbb{P}\left(\sqrt{\frac{1}{k}\sum_{j=1}^{k} y_{(j)}^2} > t + \sqrt{2eb^2 \ln\frac{e}{\gamma}}\right) \\ &\leq \exp\left(-\frac{t^2\gamma m}{5b^2}\right) \\ &\leq \exp\left(-\alpha\kappa\epsilon^2 m\right).\end{aligned}$$

It is trivial from the concentration inequality for sub-gaussian matrices that

$$\mathbb{P}(B_0^c) < 2\exp\left(-\alpha\kappa\epsilon^2 m\right).$$

Hence

$$\mathbb{P}(\tilde{\Omega}_{\epsilon,\beta}) \geq 1 - \mathbb{P}(B_0^c) - \mathbb{P}(B_3^c) \geq 1 - 3\exp\left(-\alpha\kappa\epsilon^2 m\right).$$

This completes the proof. $\square$

Theorem 3.2.2. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\mathrm{Sub}(b^2)$ and has variance 1. Let $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$ and let $y = Ax$. For $0 < \alpha < 1$ and $0 < \epsilon \leq \min\left(1, \frac{1-\sqrt{\alpha}}{40b^2\alpha\kappa}\right)$, if*

$$0 < \beta \leq \frac{(1-\sqrt{\alpha})\epsilon}{16\ln\frac{4eb^2}{(1-\sqrt{\alpha})\epsilon}} =: \eta(\epsilon, \alpha),$$

then (3.5) holds. Moreover:

$$\tilde{\beta}_{\epsilon,\alpha} \geq \frac{(1-\sqrt{\alpha})\epsilon}{32\ln\frac{1}{\epsilon}}, \quad \forall 0 < \epsilon \leq \frac{1-\sqrt{\alpha}}{4eb^2}.$$

Proof. Same as Theorem 3.1.3. $\square$

3.3 The Strong Restricted Isometry Property and the Robust Version of the Johnsen-Lindenstrauss Lemma for Small ϵ

With the estimates of $\beta_{\epsilon,\alpha}$ and $\tilde{\beta}_{\epsilon,\alpha}$, we can develop the robust version of JL Lemma and RIP properties for sub-gaussian matrices. We borrow and modify the arguments in the Gaussian case from [15, Corollary 1.2, Corollary 1.3].

Theorem 3.3.1 (Robust Johnson-Lindenstrass Lemma for small distortion rate in norms). *Suppose $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *For $\alpha \in (0,1)$ and $\epsilon \in \left(0, \min\left\{\frac{1-\sqrt{\alpha}}{4eb^2}, \frac{1-\sqrt{\alpha}}{40b^2\alpha\kappa}\right\}\right)$, let $N, m, n \in \mathbb{N}$ satisfying*

$$m > \frac{\ln \frac{3N(N-1)}{2}}{\alpha\kappa\epsilon^2}.$$

Then given any set $p_1, \ldots, p_N \in \mathbb{R}^n$ of N points:

$$(1-\epsilon)\|p_j - p_k\|_2^2 \le \frac{1}{m}\|A_T(p_j - p_k)\|_2^2 \le \frac{1}{|T|}\|A_T(p_j - p_k)\|_2^2 \le (1+\epsilon)\|p_j - p_k\|_2^2$$
$$\forall j, k \in \{1, \ldots, N\}, T \in T_{\epsilon,\alpha}.$$

holds with probability at least $1 - \frac{3N(N-1)}{2}\exp(-\alpha\kappa\epsilon^2 m)$, where $T_{\epsilon,\alpha}$ is defined as

$$T_{\epsilon,\alpha} = \left\{T \subseteq \{1, \ldots, m\} : |T^c| \le m\frac{(1-\sqrt{\alpha})\epsilon}{32\ln\frac{1}{\epsilon}}\right\}.$$

Proof. Assume $p_1, \ldots, p_N$ are pairwise disjoint. For $T \in T_{\epsilon,\alpha}$ we have

$0 \leq \beta \leq \frac{(1-\sqrt{\alpha})\epsilon}{32 \ln \frac{1}{\epsilon}}$. For $M \in \{m, |T|\}$, from Theorem 3.1.3 we have

$$\mathbb{P}\left\{\left|\frac{1}{M}\frac{\|A_T(p_j - p_k)\|_2^2}{\|p_j - p_k\|_2^2} - 1\right| \leq \epsilon, \quad \forall T \in T_{\epsilon,\alpha}\right\} \geq 1 - 3\exp(-\alpha\kappa\epsilon^2 m).$$

Then by the union bounds estimate:

$$\mathbb{P}\left\{\left|\frac{1}{M}\frac{\|A_T(p_j - p_k)\|_2^2}{\|p_j - p_k\|_2^2} - 1\right| \leq \epsilon, \quad \forall T \in T_{\epsilon,\alpha}, j \neq k\right\} \geq 1 - \frac{3N(N-1)}{2}\exp(-\alpha\kappa\epsilon^2 m) > 0.$$

provided $m > \frac{\ln \frac{3N(N-1)}{2}}{\alpha\kappa\epsilon^2}$. □

Next we are going to deduce the robust version of the restricted isometry property for sub-gaussian matrices. We'll need the following fact (see e.g. [18]):

Lemma 3.3.2. *Let $S \subseteq \{1, \ldots, n\}$ with $|S| = s$. Set*

$$\mathbb{S}_S^{n-1} = \{x \in \mathbb{S}^{n-1} : \mathrm{supp}(x) \subseteq S\}.$$

Then for any $\epsilon > 0$ there exists an $\epsilon-$net $Q_{S,\epsilon} \subseteq \mathbb{S}_S^{n-1}$ satisfying

- $\mathbb{S}_S^{n-1} \subseteq \bigcup_{q \in Q_{S,\epsilon}} B_{\frac{\epsilon}{8}}(q)$.
- $|Q_{S,\epsilon}| \leq \left(\frac{24}{\epsilon}\right)^s$.

Theorem 3.3.3 (**Strong restricted isometry property with small distortion rate in norms**). *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1*. For $\alpha \in (0,1)$ and $\epsilon \in \left(0, \min\left\{\frac{1-\sqrt{\alpha}}{4eb^2}, \frac{1-\sqrt{\alpha}}{40b^2\alpha\kappa}\right\}\right)$, let $s, m, n \in \mathbb{N}$ satisfying*

$$s \ln \frac{24en}{\epsilon s} < \alpha\kappa\epsilon^2 m - \ln 3.$$

Then

$$(1-2\epsilon)\|u\|_2^2 \leq \frac{1}{m}\|A_T u\|_2^2 \leq \frac{1}{|T|}\|A_T u\|_2^2 \leq (1+2\epsilon)\|u\|_2^2$$

holds for all $s-$sparse vector $u \in \mathbb{R}^n$ and $T \in T_{\epsilon,\alpha}$ with probability at least $1 - 3\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha\kappa\epsilon^2 m)$.

Proof. Let $M \in \{m, |T|\}$, and $S \subseteq \{1,\ldots,n\}$ with $|S| = s$. For any $\epsilon > 0$ there is an $\epsilon-$net $Q_{S,\epsilon}$. Then with probability at least $1 - 3\left(\frac{24}{\epsilon}\right)^s \exp(-\alpha\kappa\epsilon^2 m)$ we have

$$1-\epsilon \leq \frac{1}{M}\|A_T u\|_2^2 \leq 1+\epsilon, \quad \forall T \in T_{\epsilon,\alpha}, u \in Q_{S,\epsilon}.$$

Define

$$d = \sup\left\{\frac{1}{\sqrt{M}}\|A_T u\|_2 : u \in \mathbb{S}_S^{n-1}, T \in T_{\epsilon,\alpha}\right\}.$$

For any $u \in \mathbb{S}_S^{n-1}$, there is $v_u \in Q_{S,\epsilon}$ such that $\|u - v_u\|_2 \leq \frac{\epsilon}{8}$. Thus

$$\frac{1}{\sqrt{M}}\|A_T u\|_2 \leq \frac{1}{\sqrt{M}}\|A_T v_u\|_2 + \frac{1}{\sqrt{M}}\|A_T(u-v_u)\|_2 \leq \sqrt{1+\epsilon} + \frac{d\epsilon}{8}.$$

Also

$$d \leq \sqrt{1+\epsilon} + \frac{d\epsilon}{8} \quad \Longrightarrow \quad d \leq \frac{\sqrt{1+\epsilon}}{1-\frac{\epsilon}{8}} \leq \sqrt{1+2\epsilon}$$

as $\epsilon \in (0,1)$. On the other hand:

$$\frac{1}{\sqrt{M}}\|A_T u\|_2 \geq \frac{1}{\sqrt{M}}\|A_T v_u\|_2 - \frac{1}{\sqrt{M}}\|A_T(u-v_u)\|_2 \geq \sqrt{1-\epsilon} - \frac{d\epsilon}{8}.$$

So

$$d \geq \sqrt{1-\epsilon} - \frac{d\epsilon}{8} \quad \Longrightarrow \quad d \geq \frac{\sqrt{1-\epsilon}}{1-\frac{\epsilon}{8}} > \sqrt{1-2\epsilon}.$$

Hence it follows that

$$(1-2\epsilon)\|u\|_2^2 \leq \frac{1}{M}\|A_T u\|_2^2 \leq (1+2\epsilon)\|u\|_2^2, \quad \forall \operatorname{supp}(u) \subseteq S, T \in T_{\epsilon,\alpha}$$

There are in total $\binom{n}{s}$ subsets of $\{1,\ldots,n\}$ with cardinality s. Hence by the union bound argument and the Stirling's approximation $\binom{n}{s} \leq \left(\frac{en}{s}\right)^s$, it follows that

$$\begin{aligned}&\mathbb{P}\left((1-2\epsilon)\|u\|_2^2 \leq \frac{1}{m}\|A_T u\|_2^2 \leq \frac{1}{|T|}\|A_T u\|_2^2 \leq (1+2\epsilon)\|u\|_2^2, \quad \forall |\operatorname{supp}(u)| \leq s, T \in T_{\epsilon,\alpha}\right)\\ >&1-3\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha\kappa\epsilon^2 m).\end{aligned}$$

This completes the proof. □

Chapter 4

Sub-gaussian Matrices under Arbitrary Erasure of Rows with a Given Portion of Corruption

In this chapter we will further study the erasure robustness of sub-gaussian matrices for erasure ratio no more than a given number β. We will study the strong restricted isometry property (SRIP) of sub-gaussian matrices with a constant portion of corruption.

The main contributions which are provided in this chapter are establishing the following results :

- The SRIP for sub-gaussian matrices with a given portion of corruption,
- The robust version of the JL Lemma for sub-gaussian matrices under a given erasure ratio.

Fix $x \in \mathbb{S}^{n-1}$, for $\beta \in [0,1), 0 \leq \theta \leq \omega \leq \infty$, set

$$\Omega_{[\theta,\omega],\beta} = \left\{ \frac{1}{|T|}\|A_T x\|_2^2 \in [\theta,\omega] \quad \forall T \subseteq \{1,\dots,m\}, |T^c| \leq \beta m \right\},$$

$$\tilde{\Omega}_{[\theta,\omega],\beta} = \left\{ \frac{1}{m}\|A_T x\|_2^2 \in [\theta,\omega] \quad \forall T \subseteq \{1,\dots,m\}, |T^c| \leq \beta m \right\}.$$

For $\beta \in (0,1), \alpha > 0$, define

$$\theta_\beta(\alpha) = \sup\{\theta \in [0,\infty] : \mathbb{P}(\Omega_{[\theta,\infty],\beta}) \geq 1 - e^{-\alpha m} \quad \forall m \in \mathbb{N}\},$$

$$\omega_\beta(\alpha) = \inf\{\omega \in [0,\infty] : \mathbb{P}(\Omega_{[0,\omega],\beta}) \geq 1 - e^{-\alpha m} \quad \forall m \in \mathbb{N}\},$$

$$\tilde{\theta}_\beta(\alpha) = \sup\{\theta \in [0,\infty] : \mathbb{P}(\tilde{\Omega}_{[\theta,\infty],\beta}) \geq 1 - e^{-\alpha m} \quad \forall m \in \mathbb{N}\},$$

$$\tilde{\omega}_\beta(\alpha) = \inf\{\omega \in [0,\infty] : \mathbb{P}(\tilde{\Omega}_{[0,\omega],\beta}) \geq 1 - e^{-\alpha m} \quad \forall m \in \mathbb{N}\}.$$

By studying these quantities, we get the lower and upper bounds of the norm after the reduced sub-gaussian matrix acting on a vector, and thus enables us to determine the constants in the SRIP.

4.1 Estimates of $\theta_\beta(\alpha)$ and $\tilde{\theta}_\beta(\alpha)$

Recall the following facts:

1. Let $X_1, \dots, X_d$ be independent $\mathrm{Sub}(b^2)$ random variables with $\mathrm{Var}(X_i) = 1$ for all $i = 1, \dots, d$ and put $X = (X_1, \dots, X_d)$. Then it is straight forward to

see that :

$$\mathbb{E}\langle X,x\rangle^2 = \mathbb{E}\left(\sum_{j=1}^{d} x_i X_i\right)^2 = \sum_{1\le i\le j\le d} x_i x_j \mathbb{E}(X_i X_j)$$
$$= \sum_{j=1}^{d} x_j^2 \mathbb{E}(X_j^2) = \sum_{j=1}^{d} x_j^2 = \|x\|_2^2.$$

2. Given $X \sim \mathrm{Sub}(b^2)$ and $1 \le p < \infty$. From (ii) of Theorem 2.1.4 it follows that

$$\begin{aligned}
\mathbb{E}|X|^p &= p\int_0^\infty t^{p-1}\mathbb{P}(|X|\ge t)\,\mathrm{d}t \\
&\le 2p\int_0^\infty t^{p-1}e^{-\frac{t^2}{2b^2}}\,\mathrm{d}t \\
&= 2^{\frac{p}{2}}b^p p\int_0^\infty u^{\frac{p}{2}-1}e^{-u}\,\mathrm{d}u \\
&= 2^{\frac{p}{2}}b^p p\Gamma\left(\frac{p}{2}\right).
\end{aligned} \tag{4.1}$$

With the above facts, we introduce the following result which will be used to obtain the lower estimate of $\theta_\beta(\alpha)$. To prove this lemma, we use exactly the same argument as in the proof of [3, Lemma 2.2]:

Lemma 4.1.1. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\mathrm{Sub}(b^2)$ and has variance 1. Let $x \in \mathbb{S}^{n-1}$ and $Y = \|Ax\|_2^2$. Then for all $0 < q < \frac{m}{2}$ we have*

$$\mathbb{P}\left(Y \le \frac{m}{2} - q\right) < \exp\left(-\frac{q}{16b}\right).$$

Proof. For $\mu \in \mathbb{R}$, define $F(\mu) = \ln[\mathbb{E}\exp(-\mu Y)]$. By Markov's inequality:

$$\begin{aligned}
\mathbb{P}(-\mu Y \geq F(\mu) + \nu) &= \mathbb{P}(\exp(-\mu Y - F(\mu)) \geq \exp(\nu)) \\
&\leq \frac{\mathbb{E}(\exp(-\mu Y - F(\mu))}{\exp(\nu)} \\
&= \exp(-\nu)
\end{aligned}$$

for all $\nu \in \mathbb{R}$. Recall that for all $t \geq 0$ we have

$$1 - t \leq \exp(-t) \leq 1 - t + \frac{t^2}{2}.$$

Denote a_{ij} the entry of A in the $i-$th row and $j-$th colume. It follows that

$$\begin{aligned}
\mathbb{E}\exp(-\mu Y) &= \prod_{j=1}^{m} \mathbb{E}\left[\exp\left(-\mu\left|\sum_{i=1}^{n} x_i a_{ij}\right|^2\right)\right] \\
&\leq \prod_{j=1}^{m}\left[1 - \mu\mathbb{E}\left|\sum_{i=1}^{n} x_i a_{ij}\right|^2 + \frac{\mu^2}{2}\mathbb{E}\left|\sum_{i=1}^{n} x_i a_{ij}\right|^4\right] \\
&\leq \prod_{j=1}^{m}\left[1 - \mu + \frac{\mu^2}{2}(16b\Gamma(2))\right] \\
&\leq \exp\left(-m\mu + 8bm\mu^2\right).
\end{aligned}$$

Therefore

$$F(\mu) \leq -m\mu + 8bm\mu^2.$$

It follows that

$$\mathbb{P}\left(-\mu Y \geq -m\mu + 8bm\mu^2 + \nu\right) \leq \exp(-\nu).$$

Setting $\mu = \frac{1}{16b}$ and $\nu = \frac{q}{16b}$, the proof is complete. $\square$

Lemma 4.1.2 (Lower estimate of $\theta_\beta(\alpha)$). *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random* *matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *Then for* $\beta \in (0, t_{sg})$ *and* $\alpha \in \left(0, \ln\left((1-\beta)^{1-\beta}\beta^\beta\right) + \frac{q_\beta}{16b}(1-\beta)\right)$, *we have*

$$\theta_\beta(\alpha) \geq \min\left\{\frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right], \frac{1}{2} - \alpha\right\},$$

where t_{sg} *is a fixed constant in* $\left(0, \frac{1}{2}\right)$ *depending only on the sub-gaussian moment* b^2, *and* $q_\beta \in \left(0, \frac{1}{2}\right)$ *only depends on* β.

Proof. Recall that for $\theta > 0$, $\beta \in (0,1)$, $x \in \mathbb{S}^{n-1}$:

$$\Omega_{[\theta,\infty],\beta} = \left\{\frac{1}{|T|}\|A_T x\|_2^2 \geq \theta, \quad \forall T \subseteq \{1, \ldots, m\}, |T^c| \leq \beta m\right\}.$$

For $\alpha > 0$, we want to choose θ such that

$$\mathbb{P}(\Omega^c_{[\theta,\infty],\beta}) < \exp(-\alpha m).$$

Let $|T^c| = k \leq \beta m$ and $\gamma = k/m$. Then A_T is an $(m-k) \times n$ matrix with i.i.d. symmetric Bernoulli entries, therefore

$$\mathbb{P}\left(\|A_T x\|_2^2 \leq \frac{1}{2}(m-k) - q\right) \leq \exp\left(-\frac{q}{16b}\right), \quad \forall 0 < q < \frac{1}{2}(m-k).$$

Now for any $0 < q < \frac{1}{2}(m-k)$, set $q_T = \frac{q}{|T|} = \frac{q}{m-k}$. Then we have

$$\mathbb{P}\left(\|A_T x\|_2^2 \leq \left(\frac{1}{2} - q_T\right)(m-k)\right) \leq \exp\left(-\frac{q_T}{16b}(m-k)\right),$$

which is equivalent to

$$\mathbb{P}\left(\frac{1}{|T|}\|A_T x\|_2^2 \leq \frac{1}{2} - q_T\right) \leq \exp\left(-\frac{q_T}{16b}(1-\gamma)m\right). \tag{4.2}$$

Recall the Stirling's approximation:

$$\sqrt{2\pi}m^{m+\frac{1}{2}}e^{-m} \leq m! \leq em^{m+\frac{1}{2}}e^{-m}. \tag{4.3}$$

Therefore for $k \geq 1$:

$$\begin{aligned}
\binom{m}{k} &= \frac{m!}{(m-k)!k!} \leq \frac{em^{m+\frac{1}{2}}e^{-m}}{2\pi(m-k)^{m-k+\frac{1}{2}}e^{-(m-k)}k^{k+\frac{1}{2}}e^{-k}} \\
&= \frac{e}{2\pi}\left(\frac{m}{m-k}\right)^{m-k}\left(\frac{m}{k}\right)^{k}\left(\frac{m}{(m-k)k}\right)^{\frac{1}{2}} \\
&\leq \frac{e}{\sqrt{2\pi}}\left(\frac{m}{m-k}\right)^{m-k}\left(\frac{m}{k}\right)^{k}.
\end{aligned}$$

Now consider two cases:

- If $\lfloor \beta m \rfloor = 0$, then

$$\mathbb{P}(\Omega^c_{[\theta,\infty],\beta}) = \mathbb{P}\left(\|Ax\|^2 < \theta m\right) \leq \exp\left(-\left(\frac{1}{2}-\theta\right)m\right),$$

  for all $\theta \in \left(0, \frac{1}{2}\right)$. Then simply choose $\theta \leq \frac{1}{2} - \alpha$ we will have $\mathbb{P}(\Omega^c_{[\theta,\infty],\beta}) < \exp(-\alpha m)$.

- If $\lfloor \beta m \rfloor \geq 1$: observe that

$$\mathbb{P}(\Omega^c_{[\theta,\infty],\beta}) = \mathbb{P}\left(\min_{|T^c|=\beta m}\frac{1}{|T|}\|A_T x\|_2^2 < \frac{1}{2} - q_\beta\right).$$

Without loss of generality assume $\beta m =: k \in \mathbb{N}$. Then we have

$$\begin{aligned}\binom{m}{k} &\le \frac{e}{\sqrt{2\pi}}\left(\frac{m}{m-k}\right)^{m-k}\left(\frac{m}{k}\right)^{k}\\ &= \left(\frac{e}{\sqrt{2\pi}}\right)\left(\frac{1}{1-\beta}\right)^{(1-\beta)m}\left(\frac{1}{\beta}\right)^{\beta m}\\ &= \left(\frac{e}{\sqrt{2\pi}}\right)\exp\left(m\ln\left((1-\beta)^{-(1-\beta)}\beta^{-\beta}\right)\right).\end{aligned}$$

As (4.2) holds for all $|T^c| = k$, it follows that for any $q \in \left(0, \frac{1}{2}\right)$:

$$\begin{aligned}&\mathbb{P}\left(\min_{|T^c|=k}\frac{1}{|T|}\|A_T x\|_2^2 \le \frac{1}{2} - q\right)\\ \le&\binom{m}{k}\exp\left(-\frac{q}{16b}(1-\beta)m\right)\\ \le&\left(\frac{e}{\sqrt{2\pi}}\right)\exp\left[m\ln\left((1-\beta)^{-(1-\beta)}\beta^{-\beta}\right) - \frac{q}{16b}(1-\beta)m\right].\end{aligned}$$

Now we want to bound $\mathbb{P}(\Omega^c_{[\theta,\infty],\beta})$ by $C\exp(-\alpha m)$ where C is an absolute constant independent of m. Then it suffices to have $C = \frac{e}{\sqrt{2\pi}}$ and

$$\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right) + \frac{q}{16b}(1-\beta) \ge \alpha. \tag{4.4}$$

For $q \in \left(0, \frac{1}{2}\right)$, define

$$\begin{aligned}f_q(t) &= \ln\left((1-t)^{1-t}t^{t}\right) + \frac{q}{16b}(1-t)\\ &= (1-t)\ln(1-t) + t\ln(t) + \frac{q}{16b}(1-t)\end{aligned} \tag{4.5}$$

Then

$$f_q'(t) = -\ln(1-t) + \ln t - \frac{q}{16b}.$$

It is easy to see that $f_q'(t) < 0$ whenever $t < \frac{\exp\left(\frac{q}{16b}\right)}{1+\exp\left(\frac{q}{16b}\right)}$. Recall that for a sub-gaussian random variable, its sub-gaussian moment is greater than or equal to its variance, thus $b \geq 1$. Moreover, as $q \in (0, \frac{1}{2})$, we conclude that $\frac{\exp\left(\frac{q}{16b}\right)}{1+\exp\left(\frac{q}{16b}\right)} \geq \frac{1}{2}$. It follows that $f_q(t)$ is strictly decreasing on $\left(0, \frac{1}{2}\right)$, and

$$f_q\left(\frac{1}{2}\right) = \ln\frac{1}{2} + \frac{q}{32b} \leq -\ln 2 + \frac{1}{64} < 0.$$

On the other hand, note that

$$\lim_{t \to 0^+} f_q(t) = q > 0.$$

It follows that f_q has a unique root on $\left(0, \frac{1}{2}\right)$, and call this root t_q. Since we require the left-hand side of (4.4) to be strictly positive, this forces that

$\beta < t_q$.

We now analyze the behaviour of t_q as q varies. It is easy to see that

$$\lim_{q \to 0^+} t_q = 0.$$

Now consider

$$(1 - t_q)\ln(1 - t_q) + t_q \ln(t_q) + \frac{q}{16b}(1 - t_q) = 0.$$

If we choose $q' \geq q$, then we will have

$$(1 - t_q)\ln(1 - t_q) + t_q \ln(t_q) + \frac{q'}{16b}(1 - t_q) \geq 0,$$

and as $f_{q'}(t)$ is strictly decreasing on $\left(0, \frac{1}{2}\right)$, it follows that $t_{q'} \geq t_q$, i.e. t_q

increases as q increases, and

$$\lim_{q\to(\frac{1}{2})^-} t_q =: t_{sg} < (1+e^{\frac{1}{2}})^{-1}. \tag{4.6}$$

Thus $\lim_{q\to(\frac{1}{2})^-} f_q(t_{sg}) = 0$. As $f_q(t)$ is continuous in both t, q, it follows from (4.5) and (4.6) that

$$(1-t_{sg})\ln(1-t_{sg}) + t_{sg}\ln(t_{sg}) + \frac{1}{32b}(1-t_{sg}) = 0.$$

On the other hand, as $f_q(t)$ is continuous in both variables (t,q), it follows that for any $0 < t < t_{sg}$, there is $0 < q < \frac{1}{2}$ such that $f_q(t) = 0$, and therefore $f_{q'}(t) > 0$ for all $q < q' < \frac{1}{2}$.

Thus for any $\beta \in (0, t_{sg})$, there is $0 < q_\beta < \frac{1}{2}$ such that

$$(1-\beta)\ln(1-\beta) + \beta\ln(\beta) + \frac{q_\beta}{16b}(1-\beta) > 0.$$

So we can choose α in between such that

$$q_\beta \geq 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right].$$

Therefore, if $\theta \leq \frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right]$, we have

$$\mathbb{P}(\Omega^c_{[\theta,\infty],\beta}) \leq \frac{e}{\sqrt{2\pi}}\exp(-\alpha m).$$

So by the two cases above, we have

$$\alpha < \min\left\{\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right) + \frac{q_\beta}{16b}(1-\beta), \frac{1}{2}\right\},$$

and

$$\theta \leq \min\left\{\frac{1}{2} - (1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)\right], \frac{1}{2} - \alpha\right\}.$$

As f_{q_β} is decreasing on $(0, \frac{1}{2})$ and $\lim_{t\to 0^+} f_{q_\beta}(t) = q_\beta < \frac{1}{2}$, it follows that

$$\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right) + q_\beta(1-\beta) < \frac{1}{2}, \quad \forall 0 < \beta < t_{sg} < \frac{1}{2}.$$

Thus $\alpha < \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right) + \frac{q_\beta}{16b}(1-\beta)$. $\square$

For the upper estimate of $\theta_\beta(\alpha)$, we use the order statistics arguments. We slightly modify the argument from the proof of [15, Lemma 4.1].

Lemma 4.1.3 (Upper estimate of $\theta_\beta(\alpha)$). *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *For $\beta \in (0,1)$ and $\alpha \in (0,1)$, we have $\theta_\beta(\alpha) \leq 1$.*

Proof. Fix $x \in \mathbb{S}^{n-1}$, let $y = Ax$. Set $\gamma = \lfloor \beta m \rfloor$, $k = \gamma m \in \mathbb{N}$. Observe that

$$\mathbb{P}(\Omega_{[\theta,\infty],\beta}) = \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} \geq \sqrt{\theta}\right).$$

Also note that

$$\mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} \leq \sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} \mathbb{E}y_{(j)}^2} \leq \sqrt{\frac{1}{m}\sum_{j=1}^{m} \mathbb{E}y_{(j)}^2} = 1.$$

Without loss of generality assume m is large enough and $\beta \in \mathbb{Q}$ (as the general result follows from the fact that $\mathbb{Q}$ is dense in $\mathbb{R}$) such that $\beta m \in \mathbb{N}$, i.e. $\beta = \gamma$. Then if $\delta = \sqrt{\theta} - 1 > 0$, we have

$$\begin{aligned}
\mathbb{P}(\Omega_{[\theta,\infty],\beta}) &= \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} - \mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} \geq \sqrt{\theta} - \mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2}\right) \\
&\leq \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} - \mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} \geq \sqrt{\theta} - 1\right) \\
&\leq \exp\left(-\frac{\delta^2}{5b^2}(m-k)\right) = \exp\left(-\frac{\delta^2}{5b^2}(1-\beta)m\right).
\end{aligned}$$

So if $\mathbb{P}(\Omega_{[\theta,\infty],\beta}) \geq 1 - e^{-\alpha m}$ we will have $1 - e^{-\alpha m} \leq \exp\left(-\frac{\delta^2}{5b^2}(1-\beta)m\right)$. By letting $m \to \infty$, we have $1 \leq 0$, which is impossible. Hence we must have $\delta < 0$, i.e. $\theta \leq 1$, and thus $\theta_\beta(\alpha) \leq 1$. $\square$

Combining Lemma 4.1.2 and 4.1.3 we obtain the estimates of $\theta_\beta(\alpha)$ for sub-gaussian matrices:

Theorem 4.1.4. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *For $\beta \in (0, t_{sg})$ and $\alpha \in \left(0, \ln\left((1-\beta)^{1-\beta}\beta^\beta\right) + \frac{q_\beta}{16b}(1-\beta)\right)$ we have*

$$\min\left\{\frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right], \frac{1}{2} - \alpha\right\} \leq \theta_\beta(\alpha) \leq 1,$$

where t_{sg} and q_β are defined in Lemma 4.1.2.

Arguments in the estimates of $\theta_\beta(\alpha)$ can be applied to estimate the bounds of $\tilde{\theta}_\beta(\alpha)$.

Theorem 4.1.5 (Estimating $\tilde{\theta}_\beta(\alpha)$). *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1.

For $\beta \in (0, t_{sg})$ *and* $\alpha \in \left(0, \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right) + \frac{q_\beta}{16b}(1-\beta)\right)$*, we have*

$$(1-\beta)\min\left\{\frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)\right], \frac{1}{2} - \alpha\right\} \le \tilde{\theta}_\beta(\alpha) \le 1-\beta,$$

where t_{sg} *and* q_β *are defined in Lemma 4.1.2.*

Proof. Fix $x \in \mathbb{S}^{n-1}$, let $y = Ax$. Set $k = \lfloor \beta m \rfloor$ and $\gamma = k/m$. Since

$$\frac{1}{m}\sum_{j=k+1}^{m} y_{(j)}^2 = \frac{m-k}{m}\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2 = \frac{1-\gamma}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2,$$

it follows that for $\theta \ge 0$:

$$\begin{aligned}
\mathbb{P}(\tilde{\Omega}_{[\theta,\infty],\beta}) &= \mathbb{P}\left(\min_{T\in T_\gamma}\frac{1}{m}\|A_T x\|_2^2 \ge \theta\right) \\
&= \mathbb{P}\left(\frac{1}{m}\sum_{j=k+1}^{m} y_{(j)}^2 \ge \theta\right) \\
&= \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=k+1}^{m} y_{(j)}^2} \ge \sqrt{\frac{\theta}{1-\gamma}}\right) \\
&= \mathbb{P}(\Omega_{\left[\frac{\theta}{1-\gamma},\infty\right],\beta}) \ge \mathbb{P}(\Omega_{\left[\frac{\theta}{1-\beta},\infty\right],\beta}).
\end{aligned}$$

So for all $\theta \in [0, (1-\beta)\theta_\beta(\alpha))$, we have

$$\mathbb{P}(\tilde{\Omega}_{[\theta,\infty],\beta}) \ge 1 - \exp(-\alpha m),$$

and consequently from the estimates for $\theta_\beta(\alpha)$ we conclude

$$\tilde{\theta}_\beta(\alpha) \ge (1-\beta)\theta_\beta(\alpha) \ge (1-\beta)\min\left\{\frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)\right], \frac{1}{2} - \alpha\right\}.$$

To set up the upper estimate, we use the same argument as before: Without loss of generality assume $\beta \in \mathbb{Q}$ and thus $\beta = \gamma$, in which case we have

$$\mathbb{P}(\tilde{\Omega}_{[\theta,\infty],\beta}) = \mathbb{P}(\Omega_{[\frac{\theta}{1-\beta},\infty],\beta}).$$

Again using the same proof in the estimate of $\theta_\beta(\alpha)$ yields

$$\tilde{\theta}_\beta(\alpha) = (1-\beta)\theta_\beta(\alpha) \leq 1-\beta.$$

Hence the proof is complete. $\square$

4.2 Estimates of $\omega_\beta(\alpha)$ and $\tilde{\omega}_\beta(\alpha)$

We now give the estimates of $\tilde{\omega}_\beta(\alpha)$ and $\omega_\beta(\alpha)$. The argument is exactly the same as in the gaussian case from section 4 of [15] only up to slight modifications.

Theorem 4.2.1 (**Estimating** $\tilde{\omega}_\beta(\alpha)$)**.** *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1*. For $\beta \in (0,1)$ and $\alpha > 0$, we have*

$$1 \leq \tilde{\omega}_\beta(\alpha) \leq 1 + \sqrt{\frac{\alpha}{\kappa}},$$

where κ is the same as in Theorem 2.2.1.

Proof. For $\omega > 0$, by definition:

$$\tilde{\Omega}_{[0,\omega],\beta} = \left\{\frac{1}{m}\|Ax\|_2^2 \leq \omega\right\} = \tilde{\Omega}_{[0,\omega],0}.$$

We claim that $\tilde{\omega}_\beta(\alpha) \geq 1$. Assume not, then there exists $\omega < 1$ such that

$$\mathbb{P}(\tilde{\Omega}_{[0,\omega],\beta}) > 1 - \exp(-\alpha m).$$

On the other hand from the concentration of measure for sub-gaussian matrices, we have

$$\mathbb{P}(\tilde{\Omega}_{[0,\omega],\beta}) < \exp\left(-\kappa(1-\omega)^2 m\right).$$

Hence $1 - \exp(-\alpha m) < \exp\left(-\kappa(1-\omega)^2 m\right)$. Taking $m \to \infty$, we have $1 < 0$, which is a contradiction. Therefore $\tilde{\omega}_\beta(\alpha) \geq 1$.

By letting $\varepsilon = \sqrt{\frac{\alpha}{\kappa}}$, we have

$$\mathbb{P}\left(\frac{1}{m}\|Ax\|_2^2 \leq 1 + \varepsilon\right) \geq 1 - \exp(-\kappa\varepsilon^2 m) = 1 - \exp(-\alpha m).$$

Thus

$$\tilde{\omega}_\beta(\alpha) \leq 1 + \sqrt{\frac{\alpha}{\kappa}}.$$

This completes the proof. $\square$

Theorem 4.2.2 (Estimating $\omega_\beta(\alpha)$). *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *For $\beta \in (0,1)$ and $\alpha > 0$, we have*

$$\left(\sqrt{\frac{5b^2\alpha}{1-\beta}} + \sqrt{2eb^2 \ln \frac{e}{1-\beta}}\right)^2 \geq \omega_\beta(\alpha) \geq \begin{cases} 1, & \beta m < 1 \\ \frac{1}{1-\frac{\beta}{2}}, & \beta m \geq 1. \end{cases}$$

Proof. Fix $x \in \mathbb{S}^{n-1}$, let $y = Ax$. Set $k = \lfloor \beta m \rfloor$ and $\gamma = k/m$. First observe that for any $\omega \geq 0$ we have

$$\mathbb{P}(\Omega_{[0,\omega],\beta}) = \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2} \leq \sqrt{\omega}\right).$$

Also recall from Corollary 2.3.1 we have

$$\mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2} \leq \sqrt{2eb^2 \ln \frac{em}{m-k}} = \sqrt{2eb^2 \ln \frac{e}{1-\gamma}}.$$

Therefore

$$\begin{aligned}
\mathbb{P}(\Omega_{[0,\omega],\beta}) &= \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2} - \mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2} \leq \sqrt{\omega} - \mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2}\right) \\
&\geq \mathbb{P}\left(\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2} - \mathbb{E}\sqrt{\frac{1}{m-k}\sum_{j=1}^{m-k} y_{(j)}^2} \leq \sqrt{\omega} - \sqrt{2eb^2 \ln \frac{e}{1-\gamma}}\right) \\
&\geq 1 - \exp\left(-\frac{\delta^2}{5b^2}(m-k)\right) \\
&\geq 1 - \exp(-\alpha m),
\end{aligned}$$

provided that

$$\delta := \sqrt{\omega} - \sqrt{2eb^2 \ln \frac{e}{1-\gamma}} \geq \sqrt{\frac{5b^2\alpha}{1-\gamma}} > 0.$$

Hence by $0 \leq \gamma \leq \beta < 1$, it follows that $\mathbb{P}(\Omega_{[0,\omega],\beta}) \geq 1 - \exp(-\alpha m)$ if

$$\sqrt{\omega} \geq \sqrt{\frac{5b^2\alpha}{1-\beta}} + \sqrt{2eb^2 \ln \frac{e}{1-\beta}}.$$

Therefore

$$\omega_\beta(\alpha) \leq \left(\sqrt{\frac{5b^2\alpha}{1-\beta}} + \sqrt{2eb^2 \ln \frac{e}{1-\beta}}\right)^2.$$

For the lower estimate: if $k = 0$, then we have $\Omega_{[0,\omega],\beta} = \tilde{\Omega}_{[0,\omega],\beta}$, so $\omega_\beta(\alpha) \geq 1$.

If $k > 0$, then $\gamma > \frac{\beta}{2}$, and

$$\frac{1}{m-k}\sum_{j=1}^{k} y_{(j)}^2 = \frac{m}{m-k}\frac{1}{m}\sum_{j=1}^{k} y_{(j)}^2 = \frac{1}{1-\gamma}\frac{1}{m}\sum_{j=1}^{k} y_{(j)}^2.$$

It follows that for $\omega \geq 0$:

$$\begin{aligned}
\mathbb{P}(\Omega_{[0,\omega],\beta}) &= \mathbb{P}\left(\min_{T\in T_\gamma} \frac{1}{|T|}\|A_T x\|_2^2 \leq \omega\right) \\
&= \mathbb{P}\left(\frac{1}{m-k}\sum_{j=1}^{k} y_{(j)}^2 \leq \omega\right) \\
&= \mathbb{P}\left(\sqrt{\frac{1}{m}\sum_{j=1}^{k} y_{(j)}^2} \leq \sqrt{\omega(1-\gamma)}\right) \\
&= \mathbb{P}(\tilde{\Omega}_{[0,\omega(1-\gamma)],\beta}) \leq \mathbb{P}(\tilde{\Omega}_{\left[0,\omega\left(1-\frac{\beta}{2}\right)\right],\beta}).
\end{aligned}$$

Hence

$$\begin{aligned}
\omega_\beta(\alpha) &= \inf\{\omega : \mathbb{P}(\Omega_{[0,\omega],\beta}) > 1 - \exp(-\alpha m)\} \\
&\geq \inf\{\omega : \mathbb{P}(\tilde{\Omega}_{\left[0,\omega\left(1-\frac{\beta}{2}\right)\right],\beta}) > 1 - \exp(-\alpha m)\} \\
&= \frac{\tilde{\omega}_\beta(\alpha)}{1-\frac{\beta}{2}} \geq \frac{1}{1-\frac{\beta}{2}}.
\end{aligned}$$

The proof is complete. □

4.3 Main Results

By the estimates provided in sections 4.1 and 4.2, we have the following theorem:

Theorem 4.3.1. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1*. For $\beta \in (0, t_{sg})$ and $\alpha \in \left(0, \ln\left((1-\beta)^{1-\beta}\beta^\beta\right) + \frac{q_\beta}{16b}(1-\beta)\right)$, we have*

$$\min\left\{\frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right], \frac{1}{2}-\alpha\right\} \leq \theta_\beta(\alpha) \leq 1,$$

$$(1-\beta)\min\left\{\frac{1}{2}-16b(1-\beta)^{-1}\left[\alpha-\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)\right],\frac{1}{2}-\alpha\right\}\leq\tilde{\theta}_{\beta}(\alpha)\leq 1-\beta,$$

where t_{sg} and q_β are defined in Lemma 4.1.2.

Moreover for $\beta\in(0,1)$ and $\alpha\in(0,1)$:

$$\left(\sqrt{\frac{5b^2\alpha}{1-\beta}}+\sqrt{2eb^2\ln\frac{e}{1-\beta}}\right)^2\geq\omega_\beta(\alpha)\geq\begin{cases}1, & \beta m<1\\ \frac{1}{1-\frac{\beta}{2}}, & \beta m\geq 1\end{cases},$$

$$1\leq\tilde{\omega}_\beta(\alpha)\leq 1+\sqrt{\frac{\alpha}{\kappa}},$$

where κ is the same as in Theorem 2.2.1.

We see that for $\beta\in(0,t_{sg})$, and $\alpha\in\left(0,\min\left\{\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)+\frac{q_\beta}{16b}(1-\beta),\kappa\right\}\right)$, we have $0<\tilde{\theta}_\beta(\alpha)\leq\tilde{\omega}_\beta(\alpha)<2$.

Now we are at the stage to establish the strong restricted isometry property and the robust version of the JL Lemma for sub-gaussian matrices with given erasure ratio. We borrow and modify arguments as in the proof of Corollary 1.5 and 1.6 in [15].

Theorem 4.3.2 (**Strong restricted isometry property**). *Suppose that $A\in\mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys* $\mathrm{Sub}(b^2)$ *and has variance* 1. *Let $\alpha\in\left(0,\min\left\{\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)+\frac{q_\beta}{16b}(1-\beta),\kappa\right\}\right)$ and $\beta\in(0,t_{sg})$ where κ is the same as in Theorem 2.2.1, t_{sg} and q_β are the same as in Lemma 4.1.2. Let $s,m,n\in\mathbb{N}$ and $\epsilon\in(0,1)$ be such that*

$$s\ln\frac{24en}{\epsilon s}<\alpha m-\ln 2\quad\textit{and}\quad\theta_\epsilon:=\sqrt{\tilde{\theta}}-\frac{\epsilon}{8}\sqrt{\tilde{\omega}}>0,$$

where

$$\tilde{\theta} = (1-\beta)\min\left\{\frac{1}{2} - 16b(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)\right], \frac{1}{2} - \alpha\right\},$$

$$\tilde{\omega} = 1 + \sqrt{\frac{\alpha}{\kappa}},$$

Then we have

$$\begin{aligned}&\mathbb{P}\left(\theta_\epsilon\|u\|_2^2 \leq \frac{1}{m}\|A_T u\|_2^2 \leq \tilde{\omega}(1+2\epsilon)\|u\|_2^2, \quad \forall|\operatorname{supp}(u)| \leq s, |T^c| \leq \beta m\right)\\ \geq& 1 - 2\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha m),\end{aligned}$$

$$\begin{aligned}&\mathbb{P}\left(\frac{\theta_\epsilon}{1-\beta}\|u\|_2^2 \leq \frac{1}{|T|}\|A_T u\|_2^2 \leq \omega(1+2\epsilon)\|u\|_2^2, \quad \forall|\operatorname{supp}(u)| \leq s, |T^c| \leq \beta m\right)\\ \geq& 1 - 2\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha m),\end{aligned}$$

where

$$\theta_\epsilon := \left(\sqrt{\tilde{\theta}} - \frac{\epsilon}{8}\sqrt{\tilde{\omega}}\right)^2,$$

$$\omega = \left(\sqrt{\frac{5b^2\alpha}{1-\beta}} + \sqrt{2eb^2\ln\frac{e}{1-\beta}}\right)^2.$$

Proof. To prove the theorem, we only need to slightly modify the proof of Corollary 1.6 in [15], the main idea is borrowed from the proof of Lemma 5.1 in [2]. Let $T \subseteq \{1, \ldots, m\}$ be such that $|T^c| \leq \beta m$, then with probability at least $1 - 2e^{-\alpha m}$, we have

$$\sqrt{\tilde{\theta}(1-\epsilon)}\|u\|_2 \leq \frac{1}{\sqrt{m}}\|A_T u\|_2 \leq \sqrt{\tilde{\omega}(1+\epsilon)}\|u\|_2.$$

Let S, $\mathbb{S}_S^{n-1}$ and $Q_{S,\epsilon}$ be the same as in Lemma 3.3.2. Define

$$d = \sup\left\{\frac{1}{\sqrt{m}}\|A_T u\|_2, u \in \mathbb{S}_S^{n-1}, |T^c| \leq \beta m\right\}.$$

For any $u \in \mathbb{S}_S^{n-1}$, there is $v_u \in Q_{S,\epsilon}$ such that $\|u - v_u\|_2 \leq \frac{\epsilon}{8}$. Hence

$$\frac{1}{\sqrt{m}}\|A_T u\|_2 \leq \frac{1}{\sqrt{m}}\|A_T v_u\|_2 + \frac{1}{\sqrt{m}}\|A_T u - v_u\|_2 \leq \sqrt{\tilde{\omega}(1+\epsilon)} + \frac{d\epsilon}{8}.$$

By definition of d:

$$d \leq \sqrt{\tilde{\omega}(1+\epsilon)} + \frac{d\epsilon}{8} \quad \Longrightarrow \quad d \leq \sqrt{\tilde{\omega}(1+2\epsilon)}.$$

On the other hand

$$\frac{1}{\sqrt{m}}\|A_T u\|_2 \geq \frac{1}{\sqrt{m}}\|A_T v_u\|_2 - \frac{1}{\sqrt{m}}\|A_T u - v_u\|_2 \geq \sqrt{\tilde{\theta}} - \frac{\epsilon}{8}\sqrt{\tilde{\omega}}.$$

Choose $\epsilon > 0$ small enough such that $\sqrt{\tilde{\theta}} - \frac{\epsilon}{8}\sqrt{\tilde{\omega}} > 0$, and set

$$\theta_\epsilon := \left(\sqrt{\tilde{\theta}} - \frac{\epsilon}{8}\sqrt{\tilde{\omega}}\right)^2.$$

Thus with probability at least $1 - 2\left(\frac{24}{\epsilon}\right)^s \exp(-\alpha m)$, we have

$$\sqrt{\theta_\epsilon}\|u\|_2 \leq \frac{1}{\sqrt{m}}\|A_T u\|_2 \leq \sqrt{\tilde{\omega}(1+2\epsilon)}\|u\|_2,$$

for all $u \in \mathbb{R}^n$ with $\mathrm{supp}(u) \subseteq S$ and $|T^c| \leq \beta m$.

As there are $\binom{n}{s}$ subsets of $\{1, \ldots, n\}$ with cardinality s, then by the union bound argument and Stirling's approximation $\binom{n}{s} \leq \left(\frac{en}{s}\right)^s$, we have:

$$\mathbb{P}\left(\sqrt{\theta_\epsilon}\|u\|_2 \leq \frac{1}{\sqrt{m}}\|A_T u\|_2 \leq \sqrt{\tilde{\omega}(1+2\epsilon)}\|u\|_2, \quad \forall|\operatorname{supp}(u)| \leq s, |T^c| \leq \beta m\right)$$
$$\geq 1 - 2\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha m),$$

provided $s \ln \frac{24en}{\epsilon s} < \alpha m - \ln 2$.

Moreover applying the same argument above, it is easy to see that

$$\mathbb{P}\left(\sqrt{\frac{\tilde{\theta}}{1-\beta}(1-2\epsilon)}\|u\|_2 \leq \frac{1}{\sqrt{|T|}}\|A_T u\|_2 \leq \sqrt{\omega(1+2\epsilon)}\|u\|_2, \quad \forall|\operatorname{supp}(u)| \leq s, |T^c| \leq \beta m\right)$$
$$\geq 1 - 2\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha m).$$

Thus the proof is complete. □

Theorem 4.3.3 (Robust Johnson-Lindenstrauss Lemma with a given erasure ratio). *Suppose that $A \in \mathbb{R}^{m\times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\operatorname{Sub}(b^2)$ and has variance 1. Let $\beta \in (0, t_{sg})$ and $\alpha \in \left(0, \min\left\{\ln\left((1-\beta)^{1-\beta}\beta^\beta\right) + \frac{q_\beta}{16b}(1-\beta), \kappa\right\}\right)$ where κ is the same as in Theorem 2.2.1, t_{sg} and q_β are defined as in Lemma 4.1.2. Let $N, m, n \in \mathbb{N}$ be such that*

$$m > \alpha^{-1} \ln \frac{1}{N(N-1)}.$$

For any set of N points $p_1, \ldots, p_N$ in $\mathbb{R}^n$, then

$$\mathbb{P}\left\{\tilde{\theta}\|p_j - p_k\|_2^2 \leq \frac{1}{m}\|A_T(p_j - p_k)\|_2^2 \leq \tilde{\omega}\|p_j - p_k\|_2^2, \quad \forall|T^c| \leq \beta m, 1 \leq j,k \leq N, j \neq k\right\}$$
$$\geq 1 - N(N-1)\exp(-\alpha m),$$

$$\mathbb{P}\left\{\frac{\tilde{\theta}}{1-\beta}\|p_j-p_k\|_2^2 \leq \frac{1}{|T|}\|A_T(p_j-p_k)\|_2^2 \leq \omega\|p_j-p_k\|_2^2, \quad \forall |T^c| \leq \beta m, 1 \leq j,k \leq N, j \neq k\right\}$$
$$\geq 1 - N(N-1)\exp(-\alpha m),$$

where $\tilde{\theta}$, $\tilde{\omega}$ and ω are given in Theorem 4.3.2.

Proof. The proof the same as that of Corollary 1.5 in [15] up to corresponding modifications. By Theorem 4.3.1, we have

$$\mathbb{P}\left\{\tilde{\theta}\|p_j-p_k\|_2^2 \leq \frac{1}{m}\|A_T(p_j-p_k)\|_2^2 \leq \tilde{\omega}\|p_j-p_k\|_2^2, \quad \forall |T^c| \leq \beta m\right\}$$
$$\geq 1 - 2\exp(-\alpha m),$$

$$\mathbb{P}\left\{\frac{\tilde{\theta}}{1-\beta}\|p_j-p_k\|_2^2 \leq \frac{1}{|T|}\|A_T(p_j-p_k)\|_2^2 \leq \omega\|p_j-p_k\|_2^2, \quad \forall |T^c| \leq \beta m\right\}$$
$$\geq 1 - 2\exp(-\alpha m),$$

for any fixed pair (j,k) with $1 \leq j,k \leq N$ and $j \neq k$.

As there are $\binom{N}{2} = \frac{N(N-1)}{2}$ pairs (p_j, p_k) with $j \neq k$, then by the union bound argument and the assumption $m > \alpha^{-1} \ln \frac{1}{N(N-1)}$, the result follows. □

4.4 Special Case: Bernoulli Matrices

Among all random matrices with sub-gaussian entries, the Gaussian and Bernoulli matrices are of particular interest in the study of compressed sensing, as in many cases the projection matrices are drawn from one of these two distributions. It was proved in [23] and [15] that a Gaussian matrix satisfies the SRIP with any erasure ratio. This means that given an $m \times n$ Gaussian matrix and any $\beta \in (0,1)$ with $\beta m \in \mathbb{N}$, if we arbitrarily erase βm rows, the reduced matrix will still satisfy the RIP with high probability. However, this is not true in general for arbitrary sub-gaussian matrices. Normally there might be an upper bound of the erasure ratio β. For example, in the case of a Bernoulli matrix, it was seen in [23] that the erasure ratio β cannot reach $\frac{1}{2}$. The maximum possible erasure ratio in the Bernoulli case is still unknown.

Recall from Theorem 4.3.2 that if the erasure ratio is smaller than the number t_{sg} which is defined in Lemma 4.1.2, then the reduced matrix will still have the RIP with high probability. The question is that t_{sg} may not be the maximal erasure ratio, and in cases of general sub-gaussian matrices it is difficult to do further analysis at this stage. Nevertheless, in the case of a Bernoulli matrix, we may further improve the number t_{sg}. First let's recall the following well-known result related to Bernoulli random variables:

Theorem 4.4.1 (Khinchine's inequality). *Let $X_1, \ldots, X_n$ be independent symmetric Bernoulli random variables. Then for $0 < p < \infty$, there exist $a(p), b(p) > 0$ depending only on p so that*

$$a(p)\|c\|_2 \leq \left(\mathbb{E}\left|\sum_{j=1}^{n} c_j X_j\right|^p\right)^{\frac{1}{p}} \leq b(p)\|c\|_2, \quad \forall c = (c_1, \ldots, c_n) \in \mathbb{C}^n.$$

$a(p)$ and $b(p)$ are called Khinchine constants.

With Khinchine's inequality, let's see why we may expect better results in the Bernoulli case. Recall that in section 1 of this chapter we have proved Lemma 4.1.1: Suppose that $A \in \mathbb{R}^{m \times n}$ is a random matrix with i.i.d. entries such that each entry obeys $\mathrm{Sub}(b^2)$ and has variance 1. Let $x \in \mathbb{S}^{n-1}$ and $Y = \|Ax\|_2^2$. Then for all $0 < q < \frac{m}{2}$ we have

$$\mathbb{P}\left(Y \leq \frac{m}{2} - q\right) < \exp\left(-\frac{q}{16b}\right).$$

Then we use this lemma to estimate the lower bound constant in the SRIP. In the proof of Lemma 4.1.1, we need to estimate $\mathbb{E}\left|\sum_{i=1}^{n} x_i a_{ij}\right|^4$, where a_{ij} denotes the entry in the $i-$th row and $j-$th column of the matrix A. We apply (4.1) to estimate the fourth moment of the Sub(1) random variable $\sum_{i=1}^{n} x_i a_{ij}$ and obtain

$$\mathbb{E}\left|\sum_{i=1}^{n} x_i a_{ij}\right|^4 \leq 2^{\frac{4}{2}}(1^4)2\Gamma\left(\frac{4}{2}\right) = 16.$$

The problem is that the estimate given in (4.1) may not be sharp, therefore the result we obtain might be weaker than what it should be. Fortunately, in the case of Bernoulli random matrices, the Khinchine's inequality provides optimal estimates for moments of linear combinations of Bernoulli random variables. The best Khinchine constants are given by Haagerup in [14] as follows:

Theorem 4.4.2. *The best constants $a(p)$ and $b(p)$ in Khinchine's inequality are given by*

$$b(p) = \begin{cases} 1, & 0 < p \leq 2, \\ \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}\right)^{\frac{1}{p}}, & p > 2, \end{cases}$$

and

$$a(p) = \begin{cases} 2^{\frac{1}{2}-\frac{1}{p}}, & 0 < p \leq p_0, \\ \sqrt{2}\left(\frac{\Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi}}\right)^{\frac{1}{p}}, & p_0 < p < 2, \\ 1, & p \geq 2, \end{cases}$$

where $p_0 \in (0,2)$ satisfies $\Gamma\left(\frac{p_0+1}{2}\right) = \frac{\sqrt{\pi}}{2}$.

With Khinchine's inequality and the optimal Khinchine constants, and using the same argument in the proof of Lemma 4.1.1, we can improve Lemma 4.1.1. The result is as follows:

Lemma 4.4.3. *Let $A \in \mathbb{R}^{m\times n}$ be a random matrix with i.i.d. symmetric Bernoulli entries, then for any $x \in \mathbb{S}^{n-1}$ and $0 < q < 1$:*

$$\mathbb{P}\left(\|Ax\|_2^2 \le (\frac{1}{2} - q)m\right) < \exp\left(-\frac{q}{3}m\right).$$

Observe that the inequality we obtained in Lemma 4.4.3 is sharper than the one in Lemma 4.1.1. Due to the improvement from Lemma 4.1.1 to Lemma 4.4.3, we can obtain a better SRIP level in the Bernoulli case. We have the following lower estimates of $\theta_\beta(\alpha)$ and $\tilde{\theta}_\beta(\alpha)$:

Lemma 4.4.4. *Let $A \in \mathbb{R}^{m\times n}$ be a random matrix with i.i.d. symmetric Bernoulli entries. For $\beta \in (0, t_{ber})$ and $\alpha \in \left(0, \ln\left((1-\beta)^{1-\beta}\beta^\beta\right) + \frac{q_\beta}{3}(1-\beta)\right)$, we have*

$$\theta_\beta(\alpha) \ge \min\left\{\frac{1}{2} - 3(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right], 1-\alpha\right\},$$

$$\tilde{\theta}_\beta(\alpha) \ge (1-\beta)\min\left\{\frac{1}{2} - 3(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^\beta\right)\right], 1-\alpha\right\},$$

where $t_{ber} \in \left(0, \frac{1}{2}\right)$ satisfies

$$(1-t_{ber})\ln(1-t_{ber}) + t_{ber}\ln(t_{ber}) + \frac{1}{6}(1-t_{ber}) = 0.$$

The numerical solution is $t_{ber} \approx 0.0376$, and $q_\beta \in \left(0, \frac{1}{2}\right)$ is given by

$$q_\beta = \sup\{q \in (0,1) : (1-\beta)\ln(1-\beta) + \beta\ln(\beta) + \frac{q}{3}(1-\beta) > 0\}.$$

Recall that the sub-gaussian moment of a Bernoulli random variable is $b^2 = 1$, and we have $\kappa = \frac{1}{12}$ where κ is the same as in Theorem 2.2.1. Thus using the same argument as in the proof of Theorem 4.3.2, we have the strong restricted isometry property for Bernoulli matrices:

Theorem 4.4.5 (Strong restricted isometry property for Bernoulli matrices). *Let $A \in \mathbb{R}^{m\times n}$ be a random matrix with i.i.d. symmetric Bernoulli entries. Let $\beta \in (0, t_{her})$ and $\alpha \in \left(0, \min\left\{\ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right) + \frac{q_\beta}{3}(1-\beta), \frac{1}{12}\right\}\right)$ where t_{ber} and q_β are the same as in Lemma 4.4.4. Let $s, m, n \in \mathbb{N}$ and $\epsilon \in (0,1)$ be such that*

$$s \ln \frac{24en}{\epsilon s} < \alpha m - \ln 2 \quad \textit{and} \quad \theta_\epsilon := \sqrt{\tilde{\theta}} - \frac{\epsilon}{8}\sqrt{\tilde{\omega}} > 0,$$

where

$$\tilde{\theta} = (1-\beta)\min\left\{\frac{1}{2} - 3(1-\beta)^{-1}\left[\alpha - \ln\left((1-\beta)^{1-\beta}\beta^{\beta}\right)\right], \frac{1}{2} - \alpha\right\},$$

$$\tilde{\omega} = 1 + \sqrt{12\alpha},$$

Then we have

$$\mathbb{P}\left(\theta_\epsilon \|u\|_2^2 \leq \frac{1}{m}\|A_T u\|_2^2 \leq \tilde{\omega}(1+2\epsilon)\|u\|_2^2, \quad \forall |\operatorname{supp}(u)| \leq s, |T^c| \leq \beta m\right)$$
$$\geq 1 - 2\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha m),$$

$$\mathbb{P}\left(\frac{\theta_\epsilon}{1-\beta}\|u\|_2^2 \leq \frac{1}{|T|}\|A_T u\|_2^2 \leq \omega(1+2\epsilon)\|u\|_2^2, \quad \forall |\operatorname{supp}(u)| \leq s, |T^c| \leq \beta m\right)$$
$$\geq 1 - 2\left(\frac{24en}{\epsilon s}\right)^s \exp(-\alpha m),$$

where

$$\theta_\epsilon := \left(\sqrt{\tilde{\theta}} - \frac{\epsilon}{8}\sqrt{\tilde{\omega}}\right)^2,$$

$$\omega = \left(\sqrt{\frac{5\alpha}{1-\beta}} + \sqrt{2e \ln \frac{e}{1-\beta}}\right)^2.$$

Chapter 5

Gaussian Finite Frames

In this chapter we study the robustness properties of Gaussian random finite frames with fixed erasure ratio. The setting is slightly different from previous discussion, we will only consider matrices with more rows than columns, where as in CS more often we have the projection matrix with more columns. As mentioned before, the study of the finite frame theory shares several similarities with the study of CS with corruptions. In this thesis, as a starter, we focus on the Gaussian case mainly because this is the most special and important case among sub-gaussian cases. Moreover, because of the stable property of Gaussian random variables, this case is easier to deal with and we may expect more accurate estimates.

Given a random matrix $A \in \mathbb{R}^{m \times n}$ $(m > n)$ with i.i.d. standard normal entries, and given a fixed erasure ratio $\beta \in (0, \lambda)$ where $\lambda = 1 - \frac{n}{m} > 0$. For the study of finite frames, we care about the case which exactly βm rows are erased. In this chapter we will show that an $m \times n$ Gaussian frame is a NERF of level (β, C) for any $\beta \in (0, \lambda)$ and $C > 0$ depending only on β with overwhelming probability. We will fix the mistakes, revise the argument provided in [24]. We will give more accurate estimations and improve the results from [24].

5.1 The Largest Singular Value: Concentration Inequality Approach

We begin with estimating the largest singular value of a Gaussian matrix with certain portion of rows erased. We use the following well-known result:

Lemma 5.1.1. *([22, Corollary 5.35]) Let $A \in \mathbb{R}^{m\times n}(m \geq n)$ be a random matrix with i.i.d. standard normal entries. Then*

$$\mathbb{P}(s_{\max}(A) \geq \sqrt{m} + \sqrt{n} + t) \leq e^{-\frac{t^2}{2}}, \quad \forall t \geq 0.$$

Then we can prove the following estimate of the largest singular value of a Gaussian matrix:

Theorem 5.1.2 (The largest singular value). *Let $A \in \mathbb{R}^{m\times n}$ $(m > n)$ be a random matrix with i.i.d. standard normal entries. Let $\lambda = 1 - \frac{n}{m}$. Then for any $\alpha > 0$ we have*

$$\mathbb{P}\left(s_{\max}(A) \geq r(\alpha, \lambda)\sqrt{n}\right) \leq e^{-\alpha n},$$

where

$$r(\alpha, \lambda) = (1 - \lambda)^{-\frac{1}{2}} + 1 + \sqrt{2\alpha}.$$

Proof. By Lemma 5.1.1:

$$\mathbb{P}(s_{\max}(A) \geq ((1 - \lambda)^{-\frac{1}{2}} + 1)\sqrt{n} + t) \leq e^{-\frac{t^2}{2}}, \quad \forall t \geq 0.$$

Thus for any $\alpha > 0$, setting $t = \sqrt{2\alpha n}$, we have

$$\mathbb{P}(s_{\max}(A) \geq r(\alpha, \lambda)\sqrt{n}) \leq e^{-\alpha n}, \quad \forall t \geq 0,$$

where

$$r(\alpha, \lambda) = (1-\lambda)^{-\frac{1}{2}} + 1 + \sqrt{2\alpha}.$$

Thus the proof is complete. $\square$

5.2 Non-asymptotic Estimates of the Smallest Singular Values: Direct Approach

We now give the estimate of the smallest singular value of the reduced Gaussian matrix. Given an $m \times n$ ($m > n$) random matrix A with i.i.d. standard normal entries. Let $\lambda = 1 - \frac{n}{m}$ and $\beta \in (0, \lambda]$ with $\beta m \in \mathbb{N}$. We need to estimate the probability of the following event: $\{\|Ax\|_2^2 \leq t\}$ with $x \in \mathbb{S}^{n-1}$ and certain $t > 0$.

Lemma 5.2.1. *Suppose that $A \in \mathbb{R}^{m \times n}$ ($m \geq n$) is a random matrix with i.i.d. standard normal entries. Let $x \in \mathbb{S}^{n-1}$ and $\lambda = 1 - \frac{n}{m}$. Then for any $\alpha > 0$ we have*

$$\mathbb{P}(\|Ax\|_2^2 \leq q(\alpha, \lambda)n) \leq e^{-\alpha n},$$

where

$$q(\alpha, \lambda) = -(1-\lambda)^{-1} W_0(-e^{-2\alpha(1-\lambda)-\frac{6}{5}}),$$

and W_0 denotes the principal branch of the Lambert W function.

Proof. Let $Y = Ax = (y_1, \ldots, y_m)^T$. Then $\|Y\|_2^2 = \sum_{i=1}^m y_i^2$ has the χ^2 distribution with m degrees of freedom. Thus the density of $\|Y\|_2^2$ is

$$f(x) = \frac{1}{2^{m/2}\Gamma(m/2)} x^{m/2-1} e^{-x/2}.$$

It follows that for $t > 0$:

$$\begin{aligned}
\mathbb{P}(\|Ax\|_2^2 \leq t) &= \frac{1}{2^{m/2}\Gamma(m/2)} \int_0^t x^{m/2-1} e^{-x/2}\,\mathrm{d}x \\
&= \frac{1}{\Gamma(m/2)} \int_0^{\frac{t}{2}} x^{m/2-1} e^{-x}\,\mathrm{d}x \\
&= \frac{\gamma(m/2, t/2)}{\Gamma(m/2)},
\end{aligned}$$

where

$$\gamma(s, z) = \int_0^z x^{s-1} e^{-x}\,\mathrm{d}x = \sum_{i=0}^{\infty} \frac{z^{s+i} e^{-z}}{s(s+1)\dots(s+i)}$$

denotes the lower incomplete gamma function. Let $q \in (0, (1-\lambda)^{-1}]$ and $t = qn \leq m$, it follows that

$$\begin{aligned}
\gamma(m/2, t/2) &= (2/m)(t/2)^{m/2} e^{-t/2} \sum_{i=0}^{\infty} \frac{(t/2)^i}{((m-k)/2+1)\dots((m-k)/2+i)} \\
&\leq (2/m)(t/2)^{m/2} e^{-t/2} \sum_{i=0}^{\infty} \left(\frac{t/2}{m/2+1}\right)^i \\
&= (2/m)(t/2)^{m/2} e^{-t/2} \frac{1}{1-\frac{t}{m+2}} \\
&= \left(\frac{2}{m}\right)\left(\frac{m+2}{m+2-t}\right)(t/2)^{m/2} e^{-t/2} \\
&\leq 3(t/2)^{m/2} e^{-t/2},
\end{aligned}$$

where the last inequality holds whenever $(m-k) \geq 1$. Also note that we require $t \leq m$ when we estimate the infinite sum above, which forces $q \leq (1-\lambda)^{-1}$. Now we estimate $\Gamma(m/2)$. Recall the Stirling's approximation $\Gamma(z+1) \geq \sqrt{2\pi z}(z/e)^z$, thus

$$\Gamma(m/2) \geq (2/m)\sqrt{\pi m}(m/2)^{m/2}e^{-m/2}$$
$$= 2\sqrt{2\pi}(m/2)^{m/2-1/2}e^{-m/2}.$$

Hence we have

$$\begin{aligned}
\mathbb{P}(\|Ax\|_2^2 \leq t) &\leq \frac{\gamma(m/2, t/2)}{\Gamma(m/2)} \\
&\leq 3\left(\frac{t}{2}\right)^{\frac{m}{2}} e^{-\frac{t}{2}} \frac{\left(\frac{m}{2}\right)^{-\frac{m}{2}+\frac{1}{2}} e^{\frac{m}{2}}}{2\sqrt{2\pi}} \\
&\leq \frac{3}{2\sqrt{2\pi}}\left(\frac{t}{m}\right)^{\frac{m}{2}} e^{-\frac{t}{2}} e^{\frac{1}{10}m+\frac{1}{2}m} \quad (\text{as } \sqrt{x} \leq e^{\frac{x}{5}} \text{ for all } x > 0) \\
&= \frac{3}{2\sqrt{2\pi}}\left(\frac{t}{m}\right)^{\frac{m}{2}} e^{\frac{1}{2}\left(\frac{6}{5}m-t\right)} \leq (q(1-\lambda))^{m/2} \exp\left(\frac{m}{2}(-q+\frac{6}{5})\right) \\
&= \exp\left(\frac{n}{2}(1-\lambda)^{-1}\left(\ln(q(1-\lambda)) - q(1-\lambda) + \frac{6}{5}\right)\right).
\end{aligned}$$

Now consider the function

$$g(q) = \frac{1}{2}(1-\lambda)^{-1}\left(\ln(q(1-\lambda)) - q + \frac{6}{5}\right).$$

Observe that

$$\lim_{q\to 0} g(q) = -\infty \text{ and } g((1-\lambda)^{-1}) = \frac{1}{10} > 0.$$

It follows that for any $\alpha > 0$ there exists $q(\alpha, \lambda) \in (0, (1-\lambda)^{-1})$ such that $g(q(\alpha, \lambda)) = \alpha$, and the explicit form of $q(\alpha, \lambda)$ is given by

$$q(\alpha, \lambda) = -(1-\lambda)^{-1}W_0(-e^{-2\alpha(1-\lambda)-\frac{6}{5}}),$$

where W_0 denotes the principal branch of the Lambert W function. Hence for any $\alpha > 0$ we have

$$\mathbb{P}(\|Ax\|_2^2 \leq q(\alpha, \lambda)n) \leq e^{-\alpha n}.$$

Thus the proof is complete. $\square$

Remark 5.2.2. Recall that W_0 satisfies

$$W_0(z)e^{W_0(z)} = z, \quad \forall z \geq -e^{-1}.$$

Moreover, W_0 is an increasing function on $[-e^{-1}, \infty)$ with $W_0(0) = 0$ and $W_0(-e^{-1}) = -1$. Thus $-W_0(z) \in (0, 1)$ whenever $z \in (-e^{-1}, 0)$, so the number $q(\alpha, \lambda)$ in Lemma 5.2.1 automatically satisfies $q(\alpha, \lambda) \in (0, (1 - \lambda)^{-1})$.

To estimate the smallest singular value, we need to introduce the concept of $\epsilon-$nets. Recall that for $\epsilon > 0$, a subset N_ϵ of $\mathbb{S}^{n-1}$ is called an $\epsilon-$*net* if for any $x \in \mathbb{S}^{n-1}$ there exists $v \in N_\epsilon$ so that $\|x - v\|_2 \leq \epsilon$.

Lemma 5.2.3. *([22, Lemma 5.2]) For any $\epsilon > 0$, there exists an $\epsilon-$net $N_\epsilon \subseteq \mathbb{S}^{n-1}$ so that $|N_\epsilon| \leq (1 + 2\epsilon^{-1})^n$.*

The last supporting lemma we need to estimate the smallest singular value is the following result:

Lemma 5.2.4. *Let $\lambda \in (0, 1)$, $\alpha > 0$ and $\epsilon > 0$. Define*

$$\mu_\alpha(\epsilon) = \alpha + \ln(1 + 2\epsilon^{-1}),$$

$$r(\mu_\alpha(\epsilon), \lambda) = (1 - \lambda)^{-\frac{1}{2}} + 1 + \sqrt{2\mu_\alpha(\epsilon)},$$

$$q(\mu_\alpha(\epsilon), \lambda) = -(1 - \lambda)^{-1} W_0(-e^{-2\mu_\alpha(\epsilon)(1-\lambda)-\frac{6}{5}}).$$

Then for any fixed λ and α, there exists $\epsilon = \epsilon(\alpha, \lambda) > 0$ such that

$$\sqrt{q(\mu_\alpha(\epsilon), \lambda)} > \epsilon r(\mu_\alpha(\epsilon), \lambda). \tag{5.1}$$

Proof. By Remark 5.2.2 we have

$$\lim_{z\to 0} \frac{W_0(z)}{z} = \lim_{z\to 0} e^{-W_0(z)} = 1.$$

It follows that in this case as ϵ approaches 0, we have

$$q(\mu_\alpha(\epsilon), \lambda) \sim (1-\lambda)^{-1} e^{-2(\alpha+\ln(1+2\epsilon^{-1}))(1-\lambda)-\frac{6}{5}} \sim \epsilon^{2(1-\lambda)}.$$

Therefore $\sqrt{q(\mu_\alpha(\epsilon), \lambda)} \sim \epsilon^{1-\lambda}$. On the other hand it is easy to see that

$$\epsilon r(\mu_\alpha(\epsilon), \lambda) \sim \epsilon + \epsilon\sqrt{\ln(\epsilon^{-2})}.$$

As $\lambda > 0$, it follows that

$$\lim_{\epsilon\to 0^+} \frac{\epsilon^{1-\lambda}}{\epsilon + \epsilon\sqrt{\ln(\epsilon^{-2})}} = +\infty.$$

Thus in this case $\lim_{\epsilon\to 0^+} \frac{\sqrt{q(\mu_\alpha(\epsilon),\lambda)}}{\epsilon r(\mu_\alpha(\epsilon),\lambda)} = +\infty$. Hence certainly there exists $\epsilon = \epsilon(\alpha, \lambda) > 0$ such that (5.1) holds. $\square$

Now we are at the stage to prove the estimate of the smallest singular value of a Gaussian matrix. We have the following result:

Theorem 5.2.5 (The smallest singular value). *Suppose that $A \in \mathbb{R}^{m\times n}$ ($m > n$)is a random matrix with i.i.d. standard normal entries and $\lambda = 1 - \frac{n}{m} > 0$. For $\alpha > 0$, let $q(\mu_\alpha(\epsilon), \lambda)$ and $r(\mu_\alpha(\epsilon), \lambda)$ be the same as in Lemma 5.2.4. Then*

$$\mathbb{P}\left(s_{\min}(A) \le p(\alpha, \lambda)\sqrt{n}\right) \le 2e^{-\alpha n},$$

where

$$p(\alpha, \lambda) = \sup_{\epsilon \in (0,1)} \sqrt{q(\mu_\alpha(\epsilon), \lambda)} - \epsilon r(\mu_\alpha(\epsilon), \lambda) > 0.$$

Proof. For $\epsilon > 0$, let $N_\epsilon \subseteq \mathbb{S}^{n-1}$ be an $\epsilon-$net. Let $\alpha > 0$ and $v \in N_\epsilon$. Let $x \in \mathbb{S}^{n-1}$, choose $v_x \in N_\epsilon$ such that $\|x - v_x\|_2 \leq \epsilon$. It follows that

$$\|Ax\|_2 \geq \|Av_x\|_2 - \|A(x - v_x)\|_2 \geq \|Av_x\|_2 - \epsilon s_{\max}(A).$$

Let $\alpha > 0$. Then for any $p > 0$ we have

$$\begin{aligned}
&\mathbb{P}(s_{\min}(A) \leq p\sqrt{n}) \\
&\leq \sum_{v \in N_\epsilon} \mathbb{P}(\|Av\|_2 \leq p\sqrt{n} + \epsilon s_{\max}(A)) \\
&\leq \sum_{v \in N_\epsilon} \left(\mathbb{P}(\|Av\|_2 \leq p\sqrt{n} + \epsilon r(\mu_\alpha(\epsilon), \lambda)\sqrt{n}) + \mathbb{P}(s_{\max}(A) \geq r(\mu_\alpha(\epsilon), \lambda)\sqrt{n}) \right).
\end{aligned}$$

By Lemma 5.2.4 we may choose $\epsilon = \epsilon(\alpha, \lambda) > 0$ such that

$$p_\epsilon(\alpha, \lambda) := \sqrt{q(\mu_\alpha(\epsilon), \lambda)} - \epsilon r(\mu_\alpha(\epsilon), \lambda) > 0.$$

It follows from Lemma 5.2.1 that

$$\begin{aligned}
\mathbb{P}(\|Av\|_2 \leq p_\epsilon(\alpha, \lambda)\sqrt{n} + \epsilon r(\mu_\alpha(\epsilon), \lambda)\sqrt{n}) &= \mathbb{P}\left(\|Av\|_2 \leq \sqrt{q(\mu_\alpha(\epsilon), \lambda)n}\right) \\
&\leq e^{-\mu_\alpha(\epsilon)n}.
\end{aligned}$$

On the other hand by Theorem 5.1.2 we have

$$\mathbb{P}(s_{\max}(A) \geq r(\mu_\alpha(\epsilon), \lambda)\sqrt{n}) \leq e^{-\mu_\alpha(\epsilon)n}.$$

Hence by Lemma 5.2.3 we have

$$\mathbb{P}(s_{\min}(A) \leq p_\epsilon(\alpha, \lambda)\sqrt{n}) \leq \sum_{v \in N_\epsilon} 2e^{-\mu_\alpha(\epsilon)n} \leq 2(1 + 2\epsilon^{-1})^n e^{-\mu_\alpha(\epsilon)n} = 2e^{-\alpha n}.$$

By setting

$$p(\alpha, \lambda) := \sup_{\epsilon > 0} \sqrt{q(\mu_\alpha(\epsilon), \lambda)} - \epsilon r(\mu_\alpha(\epsilon), \lambda) > 0,$$

we have

$$\mathbb{P}(s_{\min}(A) \leq p(\alpha, \lambda)\sqrt{n}) \leq 2e^{-\alpha n}.$$

We are left to show that

$$p(\alpha, \lambda) = \sup_{\epsilon \in (0,1)} \sqrt{q(\mu_\alpha(\epsilon), \lambda)} - \epsilon r(\mu_\alpha(\epsilon), \lambda) > 0. \tag{5.2}$$

By Remark 5.2.2, $|W_0(z)| \leq 1$ for all $z \in [-e^{-1}, 1)$. Thus $\sqrt{q(\mu_\alpha(\epsilon), \lambda)} \leq (1 - \lambda)^{-\frac{1}{2}}$ for all $\epsilon > 0$, $\alpha > 0$ and $\lambda \in (0, 1)$. On the other hand it is easy to see that $\epsilon r(\mu_\alpha(\epsilon), \lambda) > (1 - \lambda)^{-\frac{1}{2}}$ whenever $\epsilon \geq 1$. It follows that $p_\epsilon(\alpha, \lambda) < 0$ whenever $\epsilon \geq 1$. Moreover by Lemma 5.2.4 we can choose $\epsilon \in (0, 1)$ sufficiently small such that $p_\epsilon(\alpha, \lambda) > 0$. It follows from the definition of $p(\alpha, \lambda)$ that

$$p(\alpha, \lambda) = \sup_{\epsilon \in (0,1)} p_\epsilon(\alpha, \lambda) > 0.$$

This completes the proof. □

5.3 Proof of NERF Property

Now we are at the stage to set up the main theorem for this chapter. We are going to show that a Gaussian random finite frame is a NERF with overwhelming probability.

Theorem 5.3.1. *Let $A \in \mathbb{R}^{m\times n}$ be a random matrix with i.i.d. standard normal entries. Let $\lambda = 1 - \frac{n}{m} > 0$. Then for any $\beta \in (0, \lambda)$ such that $\beta m \in \mathbb{N}$ and $\nu > 0$, A is a NERF of level $(\beta, \frac{R(\nu,\lambda,\beta)}{P(\nu,\lambda,\beta)})$ with probability at least $1 - 3e^{-\nu n}$ where*

$$P(\nu, \lambda, \beta) := \sup_{\epsilon\in(0,1)} \sqrt{q(\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon), \lambda(\beta))} - \epsilon r(\mu_{\alpha(\nu,\lambda,\beta)}, \lambda(\beta)),$$

$$R(\nu, \lambda, \beta) = (1 - \lambda(\beta))^{-\frac{1}{2}} + 1 + \sqrt{2\alpha(\nu, \lambda, \beta)},$$

$$\alpha(\nu, \lambda, \beta) = \nu + (1 - \lambda)^{-1} \ln[\beta^{-\beta}(1 - \beta)^{-(1-\beta)}],$$

$$\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon) = \alpha(\nu, \lambda, \beta) + \ln(1 + 2\epsilon^{-1}),$$

$$r(\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon), \lambda(\beta)) = (1 - \lambda(\beta))^{-\frac{1}{2}} + 1 + \sqrt{2\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon)},$$

$$q(\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon), \lambda(\beta)) = -(1 - \lambda(\beta))^{-1}W_0(-e^{-2\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon)(1-\lambda(\beta))-\frac{6}{5}}),$$

$$\lambda(\beta) = 1 - \frac{n}{(1 - \beta)m} = \frac{\lambda - \beta}{1 - \beta}.$$

All quantities above are positive.

Proof. Let $\alpha > 0$ and $T \subseteq \{1, \ldots, m\}$ with $|T^c| = \beta m$. Note that A_T is an $(m - \beta m) \times n$ matrix with i.i.d. standard normal entries. Define

$$\lambda(\beta) = 1 - \frac{n}{(1 - \beta)m} = \frac{\lambda - \beta}{1 - \beta}.$$

Thus by Theorem 5.1.2 and Theorem 5.2.5 we have

$$\mathbb{P}\left(s_{\min}(A_T) \leq p(\alpha, \lambda(\beta))\sqrt{n}\right) \leq 2e^{-\alpha n},$$

$$\mathbb{P}\left(s_{\max}(A_T) \geq r(\alpha, \lambda(\beta))\sqrt{n}\right) \leq e^{-\alpha n},$$

where

$$p(\alpha, \lambda(\beta)) := \sup_{\epsilon \in (0,1)} \sqrt{q(\mu_\alpha(\epsilon), \lambda(\beta))} - \epsilon r(\mu_\alpha(\epsilon), \lambda(\beta)),$$

$$r(\alpha, \lambda(\beta)) = (1 - \lambda(\beta))^{-\frac{1}{2}} + 1 + \sqrt{2\alpha},$$

$$\mu_\alpha(\epsilon) = \alpha + \ln(1 + 2\epsilon^{-1}),$$

$$r(\mu_\alpha(\epsilon), \lambda(\beta)) = (1 - \lambda(\beta))^{-\frac{1}{2}} + 1 + \sqrt{2\mu_\alpha(\epsilon)},$$

$$q(\mu_\alpha(\epsilon), \lambda(\beta)) = -(1 - \lambda(\beta))^{-1} W_0(-e^{-2\mu_\alpha(\epsilon)(1-\lambda(\beta))-\frac{6}{5}}).$$

There are $\binom{m}{\beta m}$ subsets $T \subseteq \{1, \ldots, m\}$ with $|T^c| = \beta m$. Use Stirling's approximation (4.3) of binomial coefficient we have

$$\binom{m}{k} \leq \frac{e}{2\pi} \left(\frac{m}{k}\right)^k \left(\frac{m}{m-k}\right)^{m-k} \left(\frac{m}{k(m-k)}\right)^{\frac{1}{2}} = \frac{e}{2\pi} \left(\frac{m}{k(m-k)}\right)^{\frac{1}{2}} [\beta^{-\beta m}(1-\beta)^{-(1-\beta)m}].$$

Note that $\frac{m}{k(m-k)} \leq 2$. Thus

$$\binom{m}{k} \leq \frac{e}{\sqrt{2\pi}} \beta^{-\beta m}(1-\beta)^{-(1-\beta)m} \leq \exp\left(n(1-\lambda)^{-1} \ln(\beta^{-\beta}(1-\beta)^{-(1-\beta)})\right).$$

Hence it follows that

$$\begin{aligned}
&\mathbb{P}\left(p(\alpha, \lambda(\beta))\sqrt{n} \leq s_{\min}(A_T) \leq s_{\max}(A_T) \leq r(\alpha, \lambda(\beta))\sqrt{n}, \forall T \subseteq \{1, \ldots, m\}, |T^c| = \beta m\right) \\
\geq &1 - 3\binom{m}{k} e^{-\alpha n} \\
\geq &1 - 3\exp(n(-\alpha + (1-\lambda)^{-1} \ln[\beta^{-\beta}(1-\beta)^{-(1-\beta)}])) \\
\geq &1 - 3\exp(-\nu n),
\end{aligned}$$

provided that $\nu = \alpha - (1-\lambda)^{-1} \ln[\beta^{-\beta}(1-\beta)^{-(1-\beta)}] > 0$. Thus by setting

$$\alpha(\nu, \lambda, \beta) = \nu + (1-\lambda)^{-1} \ln[\beta^{-\beta}(1-\beta)^{-(1-\beta)}],$$

$$\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon) = \alpha(\nu,\lambda,\beta) + \ln(1+2\epsilon^{-1}),$$

$$r(\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon), \lambda(\beta)) = (1-\lambda(\beta))^{-\frac{1}{2}} + 1 + \sqrt{2\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon)},$$

$$q(\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon), \lambda(\beta)) = -(1-\lambda(\beta))^{-1} W_0(-e^{-2\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon)(1-\lambda(\beta))-\frac{6}{5}}),$$

$$P(\nu,\lambda,\beta) := \sup_{\epsilon\in(0,1)} \sqrt{q(\mu_{\alpha(\nu,\lambda,\beta)}(\epsilon), \lambda(\beta))} - \epsilon r(\mu_{\alpha(\nu,\lambda,\beta)}, \lambda(\beta)),$$

$$R(\nu,\lambda,\beta) = (1-\lambda(\beta))^{-\frac{1}{2}} + 1 + \sqrt{2\alpha(\nu,\lambda,\beta)},$$

with probability at least $1-3e^{\nu n}$ we have

$$\max_{T\subseteq\{1,\ldots,m\},|T^c|=\beta m} Cond(A_T) \leq \frac{R(\nu,\lambda,\beta)}{P(\nu,\lambda,\beta)},$$

where $Cond(A_T)$ denotes the condition number of A_T. $\square$

Remark 5.3.2. We used two different approaches to estimate the extreme singular values:

- For the smallest singular value, we used a brute force method. We didn't use concentration inequality type argument because the concentration inequality tells that for $A \in \mathbb{R}^{m\times n}$ with $m \geq n$, we have

$$\mathbb{P}(s_{\min}(A) \leq \sqrt{m} - \sqrt{n} - t) \leq e^{-\frac{t^2}{2}}, \quad \forall t > 0.$$

  This a sharp estimate proved in [22]. We see that if we put $t = \sqrt{m} - \sqrt{n}$, then the probability is in fact zero, but the right hand side gives the bound $e^{-\frac{(\sqrt{m}-\sqrt{n})^2}{2}} > 0$. When this estimate is applied to prove the NERF property, it will give a cap on the erasure ratio. Therefore, we need a faster decay near zero.

- For the largest singular value, we take the concentration inequality for grant, this is due to the fact that this approach gives better results than the brute force method, also it is much simpler.

5.4 Numerical Examples

In this section we provide some numerical examples.

Example 5.4.1. Let $\nu = 0.1$, we will see how large the number

$$C(\nu, \lambda, \beta) := \frac{R(\nu, \lambda, \beta)}{P(\nu, \lambda, \beta)}$$

can be if we keep the ratio $\frac{\beta}{\lambda}$ is fixed.

- $\frac{\beta}{\lambda} = \frac{1}{10}$:

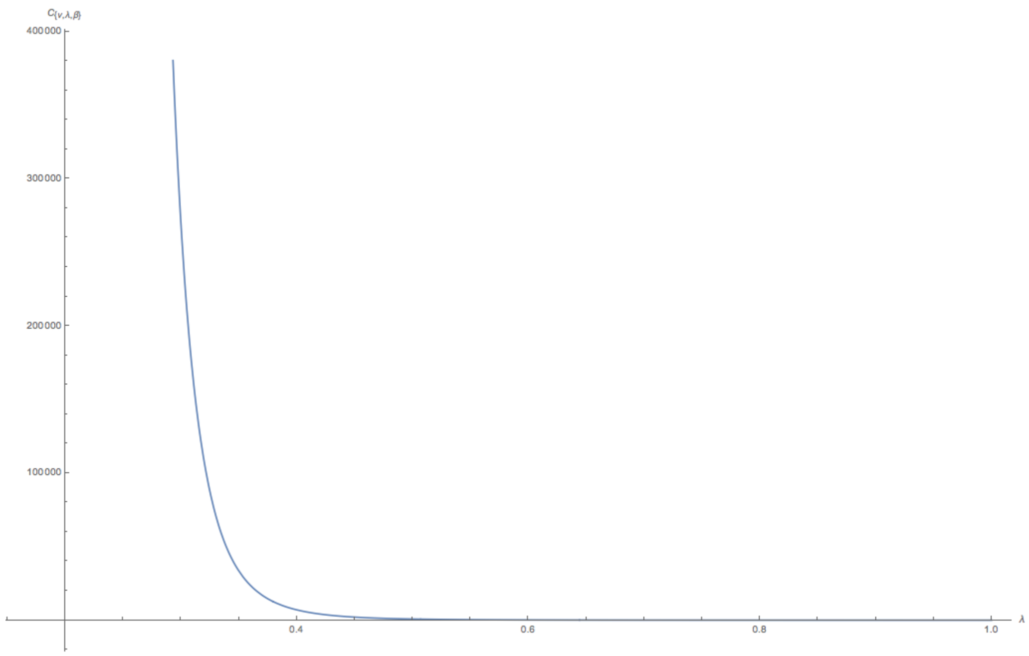

Figure 5.1: The graph of $C(\nu, \lambda, \beta)$ for $\frac{\beta}{\lambda} = \frac{1}{10}$ and $\nu = 0.1$

Table 5.1: Some numerical values of $C(\nu, \lambda, \beta)$ with $\nu = 0.1$ and $\frac{\beta}{\lambda} = \frac{1}{10}$

λ	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{4}{5}$
β	$\frac{1}{50}$	$\frac{1}{30}$	$\frac{1}{20}$	$\frac{1}{15}$	$\frac{2}{25}$
$R(\nu, \lambda, \beta)$	2.774	3.514	3.716	3.965	4.873
$P(\nu, \lambda, \beta)$	6.261×10^{-9}	4.639×10^{-5}	0.004236	0.04848	0.2028
$C(\nu, \lambda, \beta)$	4.431×10^{8}	7.575×10^{4}	877.2	81.80	24.03

We see that in this case $C(\nu, \lambda, \beta)$ decreases as λ in creases, and as λ approaches 1, $C(\nu, \lambda, \beta)$ approaches around 4.43.

- $\frac{\beta}{\lambda} = \frac{1}{2}$:

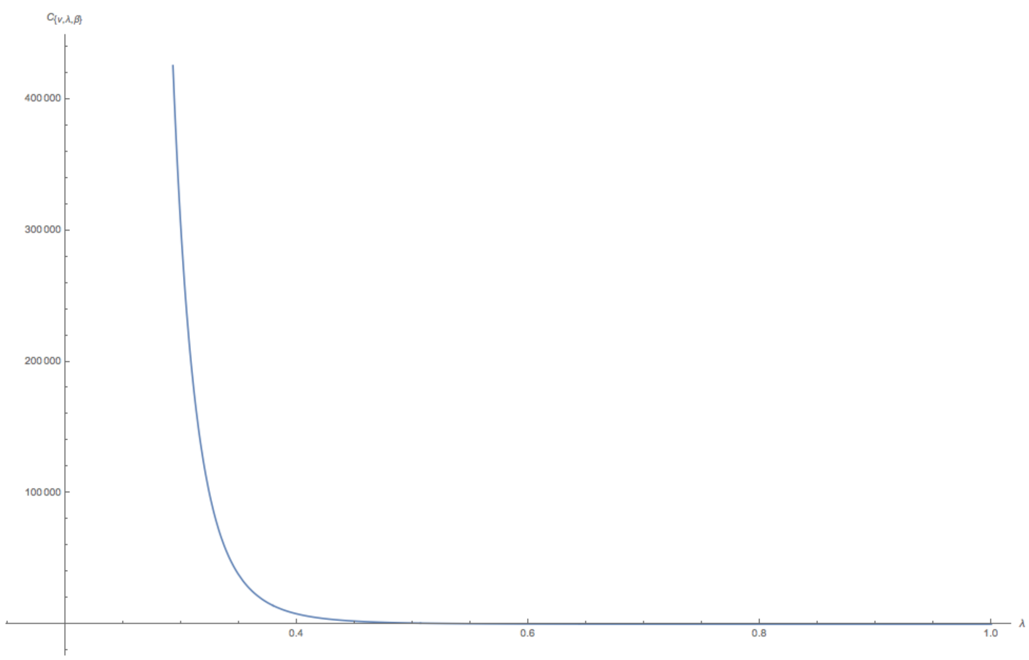

Figure 5.2: The graph of $C(\nu, \lambda, \beta)$ for $\frac{\beta}{\lambda} = \frac{1}{2}$ and $\nu = 0.1$

Table 5.2: Some numerical values of $C(\nu, \lambda, \beta)$ with $\nu = 0.1$ and $\frac{\beta}{\lambda} = \frac{1}{2}$

λ	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{4}{5}$
β	$\frac{1}{10}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{5}{12}$
$R(\nu, \lambda, \beta)$	3.067	3.364	3.790	4.419	5.364
$P(\nu, \lambda, \beta)$	6.261×10^{-9}	4.814×10^{-5}	0.004312	0.04848	0.2028
$C(\nu, \lambda, \beta)$	4.899×10^{8}	6.987×10^{4}	878.9	91.15	26.45

We see that in this case $C(\nu, \lambda, \beta)$ also decreases as λ in creases, and as λ approaches 1, $C(\nu, \lambda, \beta)$ approaches around 4.76.

- $\frac{\beta}{\lambda} = \frac{9}{10}$:

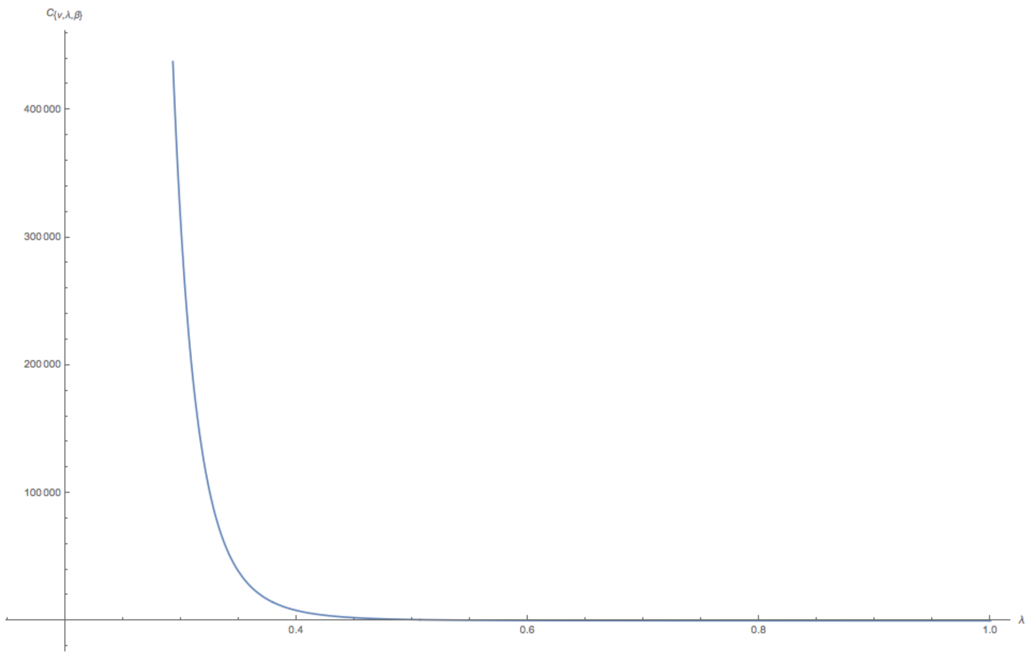

Figure 5.3: The graph of $C(\nu, \lambda, \beta)$ for $\frac{\beta}{\lambda} = \frac{1}{2}$ and $\nu = 0.1$

Table 5.3: Some numerical values of $C(\nu, \lambda, \beta)$ with $\nu = 0.1$ and $\frac{\beta}{\lambda} = \frac{9}{10}$

λ	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{4}{5}$
β	$\frac{9}{50}$	$\frac{3}{10}$	$\frac{9}{20}$	$\frac{3}{5}$	$\frac{18}{25}$
$R(\nu, \lambda, \beta)$	3.187	3.450	3.767	4.154	4.659
$P(\nu, \lambda, \beta)$	6.261×10^{-9}	4.814×10^{-5}	0.004312	0.04848	0.2028
$C(\nu, \lambda, \beta)$	5.090×10^{8}	7.167×10^{4}	873.7	85.69	22.97

In this case $C(\nu, \lambda, \beta)$ approaches around 2.83 as λ approaches 1.

Chapter 6

Summary and Discussion

In this thesis, we have studied the robustness properties of the sub-gaussian random matrices. First we have studied how large is the erasure ratio which still keeps the nearly norm preserving property for the reduced matrix. Then we've proved that sub-gaussian random matrices satisfy the strong restricted isometry property of certain level and order. In particular we take a closer look on the Bernoulli case, by employing the Khinchine's inequality we see that the result can be further improved from the general case. We also further studied the robustness properties of the Johnson-Lindenstrauss Lemma and the restricted isometry property for the sub-gaussian matrices. Last but not the least, by estimating extreme singular values of Gaussian matrices, we confirmed that Gaussian finite frames are numerically erasure robust frames. We fixed the mistakes made by the author of [24] and improved the argument.

In the future, I would like to focus on the following projects:

- We know that the Bernoulli matrices do not satisfy the SRIP with high probability if the erasure ratio is $\frac{1}{2}$ or higher. Then we would like to ask what is the maximum possible erasure ratio which still guarantees the SRIP for Bernoulli matrices.

- The explicit bounds for the condition numbers of the reduced Gaussian finite frame were not provided in Chapter 5. We are interested in whether it is possible to determine the closed form of the bounds, and if so we would like to determine whether the bounds are optimal or not. We would like to know if it is possible to further improve the existing results.

- This thesis focuses on sub-gaussian random matrices. We are also interested in what would happen if we change our settings. Would similar results hold for other types of matrices? For example, the discrete randomized Fourier transform matrices? The approach could be quite different if the settings are changed.

Bibliography

[1] Achlioptas, Dimitris. *Database-friendly random projections: Johnson-Lindenstrauss with binary coins.* Special issue on PODS 2001 (Santa Barbara, CA). J. Comput. System Sci. 66 (2003), no. 4, 671–687.

[2] Baraniuk, Richard; Davenport, Mark; DeVore, Ronald; Wakin, Michael. *A simple proof of the restricted isometry property for random matrices.* Constr. Approx. 28 (2008), no. 3, 253–263.

[3] Bennett, G.; Dor, L. E.; Goodman, V.; Johnson, W. B.; Newman, C. M. *On uncomlemented subspaces of L_p, $1 < p < 2$*, Israel J. Math. 26 (1977), no. 2, 178–187.

[4] Buldygin, V. V.; Kozacenko, Ju. V. *Sub-Gaussian random variables.* (Russian) Ukrain. Mat. Zh. 32 (1980), no. 6, 723–730.

[5] Candés, Emmanuel J.; Romberg, Justin K.; Tao, Terence. *Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information.* IEEE Trans. Inform. Theory 52 (2006), no. 2, 489–509.

[6] Candés, Emmanuel J.; Romberg, Justin K.; Tao, Terence. *Stable signal recovery from incomplete and inaccurate measurements.* Comm. Pure Appl. Math. 59 (2006), no. 8, 1207–1223.

[7] Candés, Emmanuel J.; Tao, Terence. *Decoding by linear programming*. IEEE Trans. Inform. Theory 51 (2005), no. 12, 4203–4215.

[8] Davidson, Kenneth R.; Szarek, Stanislaw J. *Local operator theory, random matrices and Banach spaces*. Handbook of the geometry of Banach spaces, Vol. I, 317–366, North-Holland, Amsterdam, 2001.

[9] DeVore, Ronald; Petrova, Guergana; Wojtaszczyk, Przemyslaw. *Instance-optimality in probability with an ℓ^1-minimization decoder*. Appl. Comput. Harmon. Anal. 27 (2009), no. 3, 275–288.

[10] Donoho, David L. *Compressed sensing*. IEEE Trans. Inform. Theory 52 (2006), no. 4, 1289–1306.

[11] Fickus, Matthew; Mixon, Dustin G. *Numerically erasure-robust frames*. Linear Algebra Appl. 437 (2012), no. 6, 1394–1407.

[12] Gordon, Y.; Litvak, A. E.; Schütt, C.; Werner, E. *On the minimum of several random variables*. Proc. Amer. Math. Soc. 134 (2006), no. 12, 3665–3675.

[13] Guédon, Olivier; Litvak, Alexander E.; Pajor, Alain; Tomczak-Jaegermann, Nicole. *Restricted isometry property for random matrices with heavy-tailed columns*. C. R. Math. Acad. Sci. Paris 352 (2014), no. 5, 431–434.

[14] Haagerup, Uffe. *The best constants in the Khintchine's inequality*, Studia Math. 70 (1981), no. 3, 231-283.

[15] Han, Bin; Xu, Zhiqiang. *Robustness properties of dimensionality reduction with gaussian random matrices*, arXiv:1501.01695v1 [cs.IT] 8 Jan 2015.

[16] Houdré, Christian; Mason, David; Rosinsky, Jan; Wellner, Jon A.. *High Dimensional Probability VI: The Banff Volume*, Birkhäuser, 2013.

[17] Ledoux, Micheal. *The concentration of measure phenomenon.* American Mathematical Society, 2001.

[18] Lorentz, G.G., von Golitschek, M., Makovoz, Yu. *Constructive Approximation: Advanced Problems.* vol. 304. Springer, Berlin (1996)

[19] Olver, F. W. J. *Uniform, exponentially improved, asymptotic expansions for the confluent hypergeometric function and other integral transforms.* SIAM J. Math. Anal. 22 (1991), no. 5, 1475–1489.

[20] Rauhut, Holger. *Compressive sensing and structured random matrices.* Theoretical foundations and numerical methods for sparse recovery, 1–92, Radon Ser. Comput. Appl. Math., 9, Walter de Gruyter, Berlin, 2010.

[21] Rudelson, Mark; Vershynin, Roman. *On sparse reconstruction from Fourier and Gaussian measurements.* Comm. Pure Appl. Math. 61 (2008), no. 8, 1025–1045.

[22] Vershynin, Roman. *Introduction to the non-asymptotic analysis of random matrices.* Compressed sensing, 210–268, Cambridge Univ. Press, Cambridge, 2012.

[23] Voroninski, Vladislav; Xu, Zhiqiang. *A strong restricted isometry property, with an application to phaseless compressed sensing.* Appl. Comput. Harmon. Anal. 40 (2016), no. 2, 386–395.

[24] Wang, Yang. *Random matrices and erasure robust frames*, arXiv: 1403.5969v1 [cs.IT] 24 Mar 2014.