Random Matrix Theory and Its Application in High-dimensional Statistics

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Danning Li

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor Of Philosophy

Tiefeng Jiang

June, 2013
Acknowledgements

Ever since I first began to study statistics, I have been fascinated with its elegant theory and broad application. Statistics is built upon a solid foundation of probability theory and can be applied to many practical areas. Especially, I am interested in the random matrix theory and the challenges and opportunities that this subject presents in statistics, physics, mathematics and many other related area. While I am writing this thesis, I have had the chance to examine many important concepts and methodologies that my research has used in a very systematic manner. I also explored some of these concepts in applied settings, which has helped me integrate my theoretical research to practical uses.

As a graduate student at the University of Minnesota, I have received tremendous supports and helps from many different people, both inside and outside the department. First of all, I would like to thank my advisor, Dr. Tiefeng Jiang, to whom I am deeply indebted for his academic support and commitment, enthusiastic guidance, and insightful contribution throughout my research. He directed me generously with his time and effort throughout my PhD study and set up a model of scholarship and professorship for me to follow in the future. As I have been studying in the School of Statistics, I also like to extend my deep thanks to the faculty members, staffs, and fellow students, whose contributions and collaborations were always helpful and illuminating to me. Finally, I should express my sincere gratitude to my friend Fan Yang for his tireless effort to review and edit my thesis.
Dedication

This thesis is dedicated to my husband, Nan Xiao, for his unconditional supports and love. He gave up his own career in China, and joined me in the U.S. to take care of me and the family. This thesis is also dedicated to my son, Luke Xiao, who becomes the greatest joy of my life. He was born during the last year of my PhD study, and accompanied with me in the whole process of writing this thesis. Finally, I also dedicate this thesis to my parents for their unwavering love and encouragement.
Abstract

This thesis mainly focuses on several classical random matrices under some special settings, which has wide applications in modern science. We study the limiting spectral distribution of the $m \times m$ upper-left corner of an $n \times n$ Haar-invariant unitary matrix, which converges to the circular law as $m \to \infty$ with $m/n \to 0$ or converges to the arc law as $m/n \to 1$. Secondly we investigate the random eigenvalues coming from the beta-Laguerre ensemble with parameter $p$, which is a generalization of the Wishart matrices of parameter $(n, p)$. In the case when the sample size $n$ is much smaller than the dimension $p$, we approximate the beta-Laguerre ensemble by a beta-Hermite ensemble which is a generalization of the Wigner matrices. As corollaries, we get that the largest and smallest eigenvalues of the complex Wishart matrix are asymptotically independent; the limiting distribution of the condition numbers; a test procedure for the spherical hypothesis test. In addition, we prove the large deviation principles for three basic statistics: the largest eigenvalue, the smallest eigenvalue and the empirical distribution of eigenvalues, and we also demonstrate that the limiting spectral distribution converges to the semicircle law as a corollary. Finally, we use a modified statistic to test the covariance structure for the $n \times p$ random matrix where $p$ is much larger than $n$. Both the law of the large number and the limiting distribution of the statistic are derived. Under the condition $\log n \ll \log p \ll n^\beta$, we also obtain the rate of convergence for the asymptotic distribution as equal to $O(\sqrt{\frac{(\log p)^3}{n}})$. Furthermore, we also use the same type of statistics to test the banded covariance structure, and show under some conditions, it holds with the similar properties.
Contents

Acknowledgements i
Dedication ii
Abstract iii
List of Tables vi
List of Figures vii

1 Background 1

2 Circular Law and Arc Law for Truncation of Random Unitary Matrix 5
  2.1 Introduction ............................................. 5
  2.2 The Arc Law and Its Proof .............................. 7
    2.2.1 The Proof of The Arc Law ......................... 8
  2.3 The Circular Law and Its Proof ........................ 19
    2.3.1 The Proof of The Circular Law .................... 20

3 Approximation of Rectangular Beta-Laguerre Ensembles and Large Deviations 26
  3.1 Introduction ............................................. 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Convergence of Laguerre Ensembles to Hermite Ensembles and Its Applications</td>
<td>30</td>
</tr>
<tr>
<td>3.3</td>
<td>Large Deviations for Eigenvalues</td>
<td>34</td>
</tr>
<tr>
<td>3.4</td>
<td>Proofs of Results in Section 3.2</td>
<td>37</td>
</tr>
<tr>
<td>3.5</td>
<td>Proofs of Results in Section 3.3</td>
<td>47</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Proof of Theorem 3.2</td>
<td>47</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Proof of Theorem 3.4</td>
<td>58</td>
</tr>
<tr>
<td>3.6</td>
<td>Appendix</td>
<td>67</td>
</tr>
<tr>
<td>4</td>
<td>Limiting Laws and Berry-Essen Bound of Statistics for Testing High Dimension Covariance Structure</td>
<td>69</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>69</td>
</tr>
<tr>
<td>4.2</td>
<td>Main Results and Applications</td>
<td>72</td>
</tr>
<tr>
<td>4.2.1</td>
<td>The Independent Case</td>
<td>72</td>
</tr>
<tr>
<td>4.2.2</td>
<td>The Dependent Case</td>
<td>73</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Testing the Covariance Structure</td>
<td>74</td>
</tr>
<tr>
<td>4.3</td>
<td>Proofs</td>
<td>75</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Proofs of Theorems 4.1, 4.2 and 4.3</td>
<td>75</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Proof of Theorem 4.4 and 4.5</td>
<td>86</td>
</tr>
<tr>
<td>4.4</td>
<td>Auxiliary Lemmas</td>
<td>96</td>
</tr>
<tr>
<td>5</td>
<td>Discussion</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>114</td>
</tr>
</tbody>
</table>
List of Tables

1.1 Classical Types of Random Matrices . . . . . . . . . . . . . . . . 3
List of Figures

2.1  Simulation for Arc Law of Three Cases: $U_{[200,185]}$, $U_{[500,477]}$ and $U_{[1000,968]}$.  

2.2  Simulation for Circular Law of Three Cases: $U_{[200,53]}$, $U_{[500,105]}$ and $U_{[1000,177]}$.  

Chapter 1

Background

Random matrix theory gained attention due to the development of quantum mechanics in 1940’s. In the late 1950’s, Wigner argued that the energy levels of a physical system can be approximated by the eigenvalues of a large random matrix. He proved that the limiting spectral distribution of a Gaussian symmetric (or Wigner) random matrix tends to the semicircular law. Since the 1960’s, Wigner, Dyson and Mehta developed the random matrix theory into a very powerful tool in mathematical physics (see Metha, 2004). Dyson introduced three types of random ensembles with elements that are real, complex and quaternion in a series of papers in 1962. Later on, Marcenko and Pastur discovered the limiting spectral property of the sample covariance matrix, which follows so-called Marcenko-Pastur law. Following these pioneers’ works, many statisticians started to work on the limiting spectral distribution for more general random matrices during 1980s (see Bai and Silverstein, 2009). Tracy and Widom first found the expression of the limiting distribution of the largest eigenvalue of a Gaussian matrix after suitable normalization in 1996. Johnstone derived the limiting distribution of the largest eigenvalues of the Wishart matrix soon after. In recent years, researches on random matrices have been focused on proving the so-called universality conjecture.

In so many years of development, one major area of random matrix research has
been centered on the study of the eigenvalues of random matrices. The other interesting direction is to investigate the entries of large dimensional random matrices. The study of large dimensional random matrices mainly includes the following topics:

- *Limiting spectral distributions of random matrices.*
- *Limiting distributions of extreme eigenvalues.*
- *Limiting distributions of the spacing.*
- *Large deviation principles for eigenvalues.*
- *Central limit theorems of linear spectral statistics.*
- *The limiting distributions of the elements of random matrices.*

More research topics of high-dimensional random matrices can be found in Anderson, Guionnet and Zeitouni (2009), Bai and Silverstein (2009) and Metha (2004).

There are so many different types of random matrices that have been investigated by physicists, mathematicians and statisticians. The classical random matrices include, among others, Gaussian symmetric matrices, Wishart matrices, Jacobi ensembles, Haar invariant matrices on compact groups. According to Dyson, these matrices can be classified based on the property of time reversal invariance, whose elements are real, complex or quaternion. These random matrices are called in different names in different areas. Table 1.1 gives a summary of these names to avoid confusion between areas. After several decades of development, these classical random matrices have been generalized by the mathematical model called beta-ensembles, see Dumitriu (2003). Many other random matrices have also attracted people’s attentions for various applications. Examples include that sample correlation matrices, metric random matrices and the adjacency and Laplace matrices of random graphs.
Table 1.1: Classical Types of Random Matrices

<table>
<thead>
<tr>
<th>Physical Name</th>
<th>Statistical Name</th>
<th>$\beta = 1$</th>
<th>$\beta = 2$</th>
<th>$\beta = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hermite (Wigner)</td>
<td>GOE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laguerre</td>
<td>Real Wishart</td>
<td>Complex W.</td>
<td>Quaternion W.</td>
<td></td>
</tr>
<tr>
<td>Jacobi</td>
<td>Real MANOVA</td>
<td>Complex M.</td>
<td>Quaternion M.</td>
<td></td>
</tr>
<tr>
<td>Haar</td>
<td>HOE</td>
<td>HUE</td>
<td>HSE</td>
<td></td>
</tr>
</tbody>
</table>

The random matrix theory has been proved to be a powerful tool in a wide variety of fields including statistics, the high-energy physics, the electrical engineering and the number theory. Random matrices are used in many physics applications such as chaotic scattering and conductance in mesoscopic systems, and statistical properties of periodically driven quantum systems. In statistics, the random matrix theory is particularly useful for analyzing high-dimensional data, which is becoming increasingly popular in many areas of scientific investigations. In these applications, the dimension $p$ can be much larger than its sample size $n$. In this setting, the classical statistical methods and results based on fixed $p$ and large $n$ are no longer applicable. Such examples of high-dimensional analysis include high-dimensional regression, hypothesis testing about high-dimensional parameters, and inference on large covariance matrices. In electrical engineering, the random matrix theory can be used to construct measurement matrices in the compressed sensing process, which is a fast developing field for providing a novel and efficient data acquisition technique that enables accurate reconstruction of highly under sampled sparse signals. For discussion of connections between random matrix theory and other fields of mathematics, particularly the number theory, one can see, e.g., Conrey, Farmer, Mezzadri and Snaith (2007).

In this thesis, we investigate several different types of random matrices under
new settings. First, we study the eigenvalues of the truncated Haar unitary matrix which has great usage in both physics and statistics. Secondly, our work is also dedicated to developing new asymptotic results for handling large dimensional datasets. In particular, we study eigenvalues of the large sample covariance matrix and develop new statistics for testing the covariance structure when \( p \) is much greater than \( n \). These results become valuable due to the rapid development of modern technology, so that the high-dimensional datasets becomes very common in various scientific and social disciplines such as climate studies, financial data analysis, information retrieval/search engines. Our results supplement the classical results in providing alternatives to analyze random matrices and high-dimensional data. The rest of this thesis is organized as follows:

- In Chapter 2, we prove that the limiting spectral distribution of the truncated Haar unitary random matrix follows either the circular law or the arc law, depending on the truncated ratio.

- In Chapter 3, we prove that for any beta-Laguerre ensemble with parameter \((p,n)\), when \( p \) is much larger than \( n \), they can be approximated by beta-Hermite ensemble. Also we derive the large deviation principle for the eigenvalues of beta-Laguerre ensemble including the sample covariance matrix as a special example. Using these two results we obtain the corresponding limiting theorems of the eigenvalues, such as the limiting spectral distributions, the limiting distributions of the extreme eigenvalues. We also show how to use condition numbers to perform a spherical test.

- In Chapter 4, we propose a new statistic for testing covariance structure in both i.i.d. case and dependent case, We derive the limiting distribution of the test statistic and its Berry-Essen bound to demonstrate the improvement over existing test statistics.

- In the final chapter, we summarize our results and conclude the thesis by introducing some questions for future studies.
Chapter 2

Circular Law and Arc Law for Truncation of Random Unitary Matrix

2.1 Introduction

Experts sometimes consider truncations of large dimensional Haar unitary matrices, which are used to describe quantum systems with absorbing boundaries (Casati, Maspero and Shepelyansky, 1999). Applications of such truncated matrices are found in optical and semiconductor superlattices (Glück, Kolovsky and Korsch, 2002), problems of quantum conductance (Forrester, 2006), distribution of resonances for open quantum maps (Fyodorov and Sommers, 2003; Schomerus and Jacquod, 2005; Nonnenmacher and Zworski, 2007; Pedrosa, Carlo, Wisniacki and Ermann, 2009).

On the other hand, truncations of Haar-invariant unitary and orthogonal matrices have applications in statistics. In particular, the singular values of such a truncation have the same distribution as those of a Jacobi matrix (Collins, 2005). In literature, a Jacobi matrix is also called a MANOVA matrix, which has been used extensively in the multivariate analysis from the field of Statistics. Some
applications of the truncations of Haar-invariant matrices can be found in, for example, Eaton (1989), Diaconis, Eaton and Lauritzen (1992) and Jiang (2009).

Certain theoretical work on truncations of Haar invariant matrices were carried out by several authors. For example, Życzkowski and Sommers (2000) derived the distributions of the absolute values of the eigenvalues and simulated the empirical distributions of the eigenvalues; Petz and Réffy (2004, 2005) analyzed the entries and proved a large deviation principle for the eigenvalues; Jiang (2006, 2009) used independent normal random variables to approximate the entries; Khoruzhenko, Sommers and Życzkowski (2010) studied the distribution of the eigenvalues of a block of an Haar orthogonal matrices, and Forrester (2010) further studied the same problem and deduced a formula for the zeros of the Kac random polynomial.

In this chapter, we shall study the spectra of a truncated block of an Haar unitary matrix. To be precise, for each $n \geq 2$, let $m = m_n < n$ be a positive integer. Let $\{U_n; n \geq 1\}$ be a sequence of Haar-invariant unitary matrices defined on the same probability space, where $U_n$ is an $n \times n$ matrix for each $n \geq 1$. Denote by $U_{[n,m]}$ the $m \times m$ upper-left corner of $U_n$, and by $\lambda_1, \cdots, \lambda_m$ the eigenvalues of $U_{[n,m]}$. Since the spectral norm of $U_{[n,m]}$ is bounded by that of $U_n$, which is one, we know that $\max_{1 \leq i \leq m} |\lambda_i| \leq 1$. Define

$$\mu_m := \frac{1}{m} \sum_{i=1}^{m} \delta_{\lambda_i}.$$  

Besides their deep theoretical results, Życzkowski and Sommers (2000) made several simulations about $\mu_m$. They found that the behavior of $\mu_m$ is decided by the ratio $m/n$. In fact, taking $n = 5$, $m = 2$ and $m = 4$, respectively, and independently generating the $m$-dimensional vector $(\lambda_1, \cdots, \lambda_m)$ many times, they observed that

"For $U_{[5,4]}$, there exist several eigenvalues close to the unit circle, while for $U_{[5,2]}$ the eigenvalues are clustered closer to the origin.”  

(2.1)

Through the study of the large deviations for the eigenvalues $\lambda_1, \cdots, \lambda_m$, Petz and Réffy (2005) obtained the following nice result: If $m/n \to \alpha \in (0,1)$, then
\(\mu_m\) converges weakly to the distribution \(\nu_0\) with the probability density function

\[
d\nu_0 = \frac{(\alpha^{-1} - 1)r}{\pi(1 - r^2)^2} \, dr \, d\varphi, \quad z = re^{i\varphi}
\]  

(2.2)
on \{z \in \mathbb{C}; |z| \leq \sqrt{\alpha}\}. The density, as a function defined on the complex plane, does not depend on the angle \(\varphi\). Therefore it is rotation-invariant.

In this chapter we will study \(\mu_m\) for the other two cases: \(\alpha = 1\) and \(\alpha = 0\). From now on, for brevity of notation, we write \(m = m_n\) if there is no confusion.

### 2.2 The Arc Law and Its Proof

In this section, we first state the arc law for truncated random Haar unitary matrix when the ratio \(\alpha = 1\), then give out the proof.

**Theorem 2.1 (Arc Law)** Assume \(\lim_{n \to \infty} m/n = 1\), then, with probability one,

\[
\mu_m := \frac{1}{m} \sum_{i=1}^{m} \delta_{\lambda_i}
\]

converges weakly to the uniform distribution on \(\{z \in \mathbb{C}; |z| = 1\}\).

Recall the first part of the observations in (2.1). Thinking that \(m = 4\) is “very close to” \(n = 5\), Theorem 2.1 says that not only “several eigenvalues close to the unit circle”, there are \(m - o(m)\) of the eigenvalues close to the circle evenly, that is, they are uniformly distributed on the unit circle in the limiting sense. See Figure 2.2.

Let \(X_n = (x_{ij})\) be an \(n \times n\) Haar-invariant orthogonal or unitary matrix, that is, \(X_n\) generates the Haar probability measure on the classical compact group \(O(n)\) or \(U(n)\). It is known that the empirical distribution of the eigenvalues of \(X_n\) converges to the arc law as \(n \to \infty\). See, for example, Diaconis and Shahshahani (1994) and Diaconis and Evans (2001) in this direction. Theorem 2.1 says that the arc law still holds if the size of a truncated block of an Haar unitary matrix is large enough.
2.2.1 The Proof of The Arc Law

Theorem 2.1 is proved by using a polynomial method. Let \( g(z) = \det(zI_m - U_{[n,m]}) = z^m + b_{m-1}z^{m-1} + \cdots + b_0 \) be the characteristic polynomial of \( U_{[n,m]} \).

We first study the coefficients \( b_i \)'s by using the symmetric functions associated with the partitions of integers, an identity from Diaconis and Evans (2001), the Selberg integral and the Jacobi ensembles. Based on the work of Erdős-Turán (1950), Granville (2007) and Hughes and Nikeghbali (2008) give the convergence of the empirical distribution of the roots of a polynomial converging to the uniform distribution on \( \{z \in \mathbb{C}; |z| = 1\} \). The two steps are then combined to complete the proof.

Let \( U_n \) be an \( n \times n \) Haar-invariant unitary matrix, that is, the entries of the unitary matrix \( U_n \) are random variables having the same joint distribution as that of \( \mathbf{V}U_n \) and that of \( U_n\mathbf{V} \) for any \( n \times n \) unitary matrix \( \mathbf{V} \).

**Lemma 2.1** *(Theorem 2.1 from Diaconis and Evans (2001))*

Consider \( a = (a_1, \ldots, a_k) \) and \( b = (b_1, \ldots, b_k) \) with \( a_j, b_j \in \{0, 1, 2, \ldots\} \). Then for \( n \geq \sum_{j=1}^{k} ja_j \lor \sum_{j=1}^{k} jb_j \),

\[
\mathbb{E} \left[ \prod_{j=1}^{k} (\text{Tr}(U_n^{a_j}))^{a_j} (\text{Tr}(U_n^{b_j}))^{b_j} \right] = \delta_{ab} \prod_{j=1}^{k} j^{a_j} a_j!
\]

(2.3)

where \( \delta_{ab} \) is Kronecker’s delta.
See also Diaconis and Shahshahani (1994) for this. The following well-known formula can be found in many places, e.g., Mehta (1991) and Forrester (2007).

**Lemma 2.2 (Selberg integral)** Let \( N \geq 2 \) be an integer, and \( \alpha, \beta \) and \( \gamma \) be positive numbers. Then

\[
\int_0^1 \int_0^1 \cdots \int_0^1 \prod_{i=1}^N x_i^{\alpha-1}(1-x_i)^{\beta-1} \cdot \prod_{1 \leq j < k \leq N} |x_j - x_k|^{2\gamma} \, dx_1 \, dx_2 \cdots dx_N
\]

\[
= \prod_{l=0}^{N-1} \frac{\Gamma(1 + \gamma + l\gamma)\Gamma(\alpha + l\gamma)\Gamma(\beta + l\gamma)}{\Gamma(1 + \gamma)\Gamma(\alpha + \beta + (N + l - 1)\gamma)}.
\]

Let \( f_d(x) = a_0 + a_1x + \cdots + a_dx^d \) where \( a_0, \ldots, a_d \) are complex numbers with \( a_0a_d \neq 0 \). Define

\[
L(f) = \frac{1}{\sqrt{|a_0a_d|}} \sum_{j=0}^d |a_j|.
\] (2.4)

The following result comes from Theorem 1.3 from Granville (2007), which is an improvement of Erdős-Turán (1950). See also Theorem 7 from Hughes and Nikeghbali (2008) in the same spirit. A recent development on the distributions of the roots of random polynomial is studied by Kabluchko and Zaporozhets (2012).

**Lemma 2.3** Suppose \( f_1, f_2, \ldots \) is a sequence of polynomials of complex coefficients where \( f_d \) has degree \( d \) and \( f_d(0) \neq 0 \). Let \( \lambda_1, \cdots, \lambda_d \) be the roots of \( f_d \) and

\[
\mu_d = \frac{1}{d} \sum_{j=1}^d \delta_{\lambda_j}
\]

be the empirical measure for \( d \geq 1 \). If \( \frac{1}{d} \log L(f_d) \to 0 \) as \( d \to \infty \), then \( \mu_d \) converges weakly to the uniform distribution on \( \{ z \in \mathbb{C}; |z| = 1 \} \).

From Lemma 2.3, it is trivial to see that the empirical distribution of the roots of a Kac polynomial (the coefficients of the polynomial are independent standard normals) converges weakly to the uniform distribution on \( \{ z \in \mathbb{C}; |z| = 1 \} \) as
the degree of the polynomial goes to infinity. The same is also true for a Littlewood polynomial (the coefficients of the polynomial are independent symmetric Bernoulli random variables).

**Lemma 2.4** Let $\Gamma(x)$ be the gamma function and $a$ be a real number. Then

$$\lim_{x \to +\infty} \frac{\Gamma(x + a)}{x^a \Gamma(x)} = 1.$$

**Proof.** By the Stirling formula (see, Gamelin (2001) or from Ahlfors (1979)),

$$\log \Gamma(z) = z \log z - z - \frac{1}{2} \log z + \log \sqrt{2\pi} + \frac{1}{12z} + O\left(\frac{1}{x^3}\right)$$

as $x = \text{Re}(z) \to +\infty$. Then

$$\log \frac{\Gamma(x + a)}{\Gamma(x)} = (x + a) \log(x + a) - x \log x - a - \frac{1}{2} (\log(x + a) - \log x) + O\left(\frac{1}{x}\right)$$

as $x \to \infty$. First, use the fact that $\log(1 + t) \sim t + O(t^2)$ as $t \to 0$ to get

$$(x + a) \log(x + a) - x \log x = a \log x + a + O\left(\frac{1}{x}\right)$$

as $x \to \infty$. Evidently, $\log(x + a) - \log x = O(1/x)$ as $x \to +\infty$. Thus, we have from (2.5)

$$\log \frac{\Gamma(x + a)}{\Gamma(x)} = a \log x + O\left(\frac{1}{x}\right)$$

as $x \to \infty$. The proof is complete.

For $n \geq 2$ and variables $x_1, \cdots, x_n$, recall the elementary symmetric function and power symmetric function as follows.

$$e_k = e_k(x_1, \cdots, x_n) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}; \quad (2.6)$$

$$p_k = p_k(x_1, \cdots, x_n) = \sum_{i=1}^n x_i^k \quad \text{and} \quad p_{\lambda} = p_{\lambda_1} p_{\lambda_2} \cdots \quad (2.7)$$

for any partition $\lambda = (\lambda_1, \lambda_2, \cdots)$ with $e_0 = p_0 = 1$. 
LEMMA 2.5 Let the characteristic polynomial of an $n \times n$ Haar invariant unitary matrix $U_n$ be denoted by

$$f_n(z) = \det(zI_n - U_n) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$$

for all $z \in \mathbb{C}$. Then $E(|a_k|^2) = 1$ for $0 \leq k \leq n - 1$.

Proof. First, let the eigenvalues of $U_n$ be $z_1, \cdots, z_n$. Then we know

$$a_{n-k} = (-1)^ke_k(z_1, \cdots, z_n) \quad (2.8)$$

for $1 \leq k \leq n$. Second, we know the basic formula between $e_k$ and $p_{\lambda}$:

$$e_k = \sum_{\lambda \vdash k, \rho \vdash k} \epsilon_{\lambda}z_1^{-1}p_{\lambda} \quad (2.9)$$

where the sum runs over all partition $\lambda = (1^{m_1}2^{m_2} \cdots)$ of $k$, $\epsilon_{\lambda} = (-1)^{k-l(\lambda)}$ with $l(\lambda) = m_1 + m_2 + \cdots$ being the length of the partition $\lambda$ and

$$z_\lambda = \prod_{j \geq 1} j^{m_j}m_j!.$$

See Proposition 7.7.6 from Stanley (2001) or (2.14') from Macdonald (1998). From (2.8) and (2.9) we see that

$$E(|a_{n-k}|^2) = E(|e_k|^2) = \sum_{\lambda \vdash k, \rho \vdash k} \epsilon_{\lambda}\epsilon_{\rho}z_\lambda^{-1}z_\rho^{-1}E(p_{\lambda}\overline{p}_{\rho}) \quad (2.10)$$

for all $1 \leq k \leq n$. Notice $p_k = p_k(z_1, \cdots, z_n) = tr(U_n^k)$. Thus, for $\lambda = (1^{m_1}2^{m_2} \cdots)$, we have

$$p_{\lambda} = p_{\lambda_1}p_{\lambda_2} \cdots = \prod_{j \geq 1} \left(\text{tr}(U_n^j)\right)^{m_j}.$$ 

It then follows from (2.3) that

$$E(|p_{\lambda}|^2) = z_\lambda \text{ and } E(p_{\lambda}\overline{p}_{\rho}) = 0 \text{ for } \lambda \neq \rho.$$
By (2.10) we have

\[ E(|a_k|^2) = \sum_{\lambda \vdash k} \frac{1}{z_\lambda} = 1 \]

for \( k = 0, 1, \cdots, n - 1 \), where the last identity holds since the conjugacy class associated with \( \lambda \) in the symmetric group \( S_k \) has \( k! / z_\lambda \) permutations, so that \( \sum_{\lambda \vdash k} k! / z_\lambda = k! \).  

**Lemma 2.6** For \( 1 \leq m \leq n \), let \( U_n \) be an \( n \times n \) Haar unitary matrix and \( U_{[n,m]} \) be the \( m \times m \) upper-left corner of \( U_n \). Set \( g(z) = \det(zI_m - U_{[n,m]}) = z^m + b_{m-1}z^{m-1} + \cdots + b_0 \). Then

\[ E(|b_{m-k}|^2) = \frac{\binom{m}{k}}{\binom{n}{k}} \]

for all \( k = 1, 2 \cdots, m \).

**Proof.** It is known that

\[ (-1)^k b_{m-k} = \sum_{1 \leq j_1 < \cdots j_k \leq m} \det(U_{j_1, \cdots, j_k}) \quad (2.11) \]

for any \( 1 \leq k \leq m \), where \( U_{j_1, \cdots, j_k} \) is the \( k \times k \) principal minor of \( U_{[n,m]} \), hence the \( k \times k \) principal minor of \( U_n \) formed by rows \( j_1, \cdots, j_k \) and columns \( j_1, \cdots, j_k \) of \( U_n \). By the Haar invariant property, exchanging any two different rows or/and exchanging two different columns do not change the joint distribution of the entries. Thus, \( \det(U_{j_1, \cdots, j_k}) \) and \( \det(U_{1, \cdots, k}) \) have the same probability distribution. It follows that

\[ E\left(|\det(U_{j_1, \cdots, j_k})|^2\right) = E\left(|\det(U_{1, \cdots, k})|^2\right) \quad (2.12) \]

for any \( 1 \leq j_1 < \cdots < j_k \leq m \). On the other hand, any two different \( k \times k \) minors \( \det(U_{j_1, \cdots, j_k}) \) and \( \det(U_{l_1, \cdots, l_k}) \) are uncorrelated, that is,

\[ E\left(\det(U_{j_1, \cdots, j_k}) \cdot \overline{\det(U_{l_1, \cdots, l_k})}\right) = 0 \quad (2.13) \]
for any $1 \leq j_1 < \cdots < j_k \leq m$ and $1 \leq l_1 < \cdots < l_k \leq m$ with \{{j_1, \cdots, j_k}\} \neq \{{l_1, \cdots, l_k}\}. In fact, suppose $j_s \notin \{l_1, \cdots, l_k\}$ for some $s \in \{1, \cdots, k\}$. For any $\theta \in \mathbb{R}$, set $A = \text{diag}(1, 1, \cdots, 1, e^{i\theta}, 1, \cdots, 1)$ where $e^{i\theta}$ appears at the $s$-th position.

Then, by the Haar invariance, $AU_n$ and $U_n$ have the same distribution. Notice the transform $AU_n$ only changes the $j_s$-th row of $U_n$ by multiplying each entry in the row with $e^{i\theta}$, and keep all of the other entries of $U_n$ the same. Consequently,

$$E \left( \det(U_{j_1, \cdots, j_k}) \cdot \det(U_{l_1, \cdots, l_k}) \right) = E \left( \det((AU_n)_{j_1, \cdots, j_k}) \cdot \det((AU_n)_{l_1, \cdots, l_k}) \right) = e^{i\theta} \cdot E \left( \det(U_{j_1, \cdots, j_k}) \cdot \det(U_{l_1, \cdots, l_k}) \right)$$

for any $\theta \in \mathbb{R}$, which implies (2.13). From (2.11), (2.12) and (2.13) we conclude

$$E |b_{m-k}|^2 = \binom{m}{k} \cdot E \left( |\det(U_{1, \cdots, k})|^2 \right) \quad (2.14)$$

for all $1 \leq k \leq m$. Now define

$$\det(zI_n - U_n) = z^n + a_{n-1}z^{n-1} + \cdots + a_0.$$ 

Taking $m = n$ in (2.14) and by Lemma 2.5 we have that

$$1 = E |a_{n-k}|^2 = \binom{n}{k} \cdot E \left( |\det(U_{1, \cdots, k})|^2 \right)$$

for all $1 \leq k \leq n$. This and (2.14) yield the desired conclusion. ■

**Lemma 2.7** Let $m = m_n$ be an integer with $1 \leq m < n$ for each $n \geq 2$. Let \{U_n; n \geq 1\} be a sequence of Haar unitary matrices defined on the same probability space, where $U_n$ is an $n \times n$ matrix for each $n \geq 2$. Denote by $W_{[n,m]}$ the $m \times m$ lower-right corner of $U_n$. If $\lim_{n \to \infty} m/n = 0$, then

$$\lim_{n \to \infty} \frac{1}{n} \log |\det(W_{[n,m]})| = 0 \quad \text{a.s.} \quad (2.15)$$

**Proof.** For any $\epsilon > 0$, we claim that

$$P \left( \frac{1}{n} \log |\det(W_{[n,m]})| > \epsilon \right) \leq e^{-n\epsilon} \quad (2.16)$$
as \( n \) is sufficiently large. If this is true, then the sum of the above probabilities over all \( n \geq 2 \) is finite. Therefore, (2.15) follows from the Borel-Cantelli lemma.

Let \( U_{[n,m]} \) the \( m \times m \) upper-left corner of \( U_n \). By the Haar invariance, \( U_{[n,m]} \) and \( W_{[n,m]} \) have the same distribution. Thus, to prove (2.16), or to prove the lemma, it suffices to show, for any \( \epsilon > 0 \),

\[
P\left( \frac{1}{n} \left| \log |\det (U_{[n,m]})| \right| > \epsilon \right) \leq e^{-n\epsilon}
\]

(2.17)
as \( n \) is sufficiently large. Now let us prove this.

Set \( V_m = U_{[n,m]} U^*_{[n,m]} \) for all \( n \). Notice \( |\det (V_m)|^{1/2} = |\det (U_{[n,m]})| \leq 1 \) since \( U_{[n,m]} \) is a contraction map, and hence all of its eigenvalues are inside the unit circle on the complex plane. Therefore,

\[
P\left( \frac{1}{n} \left| \log |\det (U_{[n,m]})| \right| > \epsilon \right) = P\left( \frac{1}{|\det (V_m)|} > e^{2n\epsilon}\right)
\]

\[
\leq e^{-2n\epsilon} E \frac{1}{|\det (V_m)|^t} = e^{-2n\epsilon} E \frac{1}{(\lambda_1 \cdots \lambda_m)^t}
\]

(2.18)
for any \( t > 0 \) by the Markov inequality, where \( 0 \leq \lambda_1 \leq \cdots \leq \lambda_m \) are the eigenvalues of \( V_m \). By (3.6) and (3.15) from Forrester (2006) or (2.5) from Jiang (2009) we know that \( \lambda_1, \cdots, \lambda_m \) actually come from the Jacobi ensemble with probability density function

\[
f(\lambda_1, \cdots, \lambda_m)
\]

\[
= \frac{1}{C(m, n)} \prod_{i=1}^{m} (1 - \lambda_i)^{n-2m} \cdot \prod_{1 \leq i < j \leq m} (\lambda_i - \lambda_j)^2 I(0 \leq \lambda_1 \leq \cdots \leq \lambda_m \leq 1)
\]

(2.19)
where \( C(m, n) > 0 \) is a constant depending on \( m \) and \( n \). Since \( f(\lambda_1, \cdots, \lambda_m) \) is a probability density function,

\[
C(m, n) = \int_{0 \leq \lambda_1 \leq \cdots \leq \lambda_m \leq 1} \prod_{i=1}^{m} (1 - \lambda_i)^{n-2m} \cdot \prod_{1 \leq i < j \leq m} (\lambda_i - \lambda_j)^2 d\lambda_1 \cdots d\lambda_m
\]

\[
= \frac{1}{m!} \int_{0}^{1} \cdots \int_{0}^{1} \prod_{i=1}^{m} (1 - \lambda_i)^{n-2m} \cdot \prod_{1 \leq i < j \leq m} (\lambda_i - \lambda_j)^2 d\lambda_1 \cdots d\lambda_m.
\]
Now, taking $\alpha = 1$, $\beta = n - 2m + 1$, $\gamma = 1$ and $N = m$ in Lemma 2.2, we obtain

$$C(m, n) = \frac{1}{m!} \prod_{l=0}^{m-1} \frac{\Gamma(l + 2)\Gamma(l + 1)\Gamma(n - 2m + l + 1)}{\Gamma(2)\Gamma(n - m + l + 1)}.$$

(2.20)

On the other hand, for fixed $t \in (0, 1)$, by (2.19),

$$E \left( \frac{1}{(\lambda_1 \cdots \lambda_m)^t} \right) = \frac{1}{C(m, n)} \int \cdots \int \prod_{i=1}^{m} \lambda_i^{-t}(1 - \lambda_i)^{n - 2m} \cdot \prod_{1 \leq i < j \leq m} (\lambda_i - \lambda_j)^2 \, d\lambda_1 \cdots d\lambda_m$$

$$= \frac{1}{m! C(m, n)} \int_0^1 \cdots \int_0^1 \prod_{i=1}^{m} \lambda_i^{-t}(1 - \lambda_i)^{n - 2m} \cdot \prod_{1 \leq i < j \leq m} (\lambda_i - \lambda_j)^2 \, d\lambda_1 \cdots d\lambda_m.$$

Take $\alpha = 1 - t$, $\beta = n - 2m + 1$, $\gamma = 1$ and $N = m$ in Lemma 2.2 to get

$$E \left( \frac{1}{(\lambda_1 \cdots \lambda_m)^t} \right) = \frac{1}{m! C(m, n)} \prod_{l=0}^{m-1} \frac{\Gamma(l + 2)\Gamma(l + t + 1)\Gamma(n - 2m + l + 1)}{\Gamma(2)\Gamma(n - m - t + l + 1)}$$

$$= \prod_{l=0}^{m-1} \frac{\Gamma(l + t + 1)}{\Gamma(l + 1)} \cdot \frac{\Gamma(n - m + l + 1)}{\Gamma(n - m + l - t + 1)}.$$

Rearranging the index $l$ above, and joint with (2.18) we conclude

$$P \left( \frac{1}{n} \log |\det(U_{[n,m]}^2)| > \epsilon \right) \leq e^{-2nt} \prod_{l=1}^{m} \frac{\Gamma(l - t)}{\Gamma(l)} \cdot \prod_{l=1}^{m} \frac{\Gamma(n - m + l)}{\Gamma(n - m + l - t)}.$$

(2.21)

Now by Lemma 2.4, there exist integer $m_0 \geq 2$ such that

$$\frac{1}{2} x^{-t} \leq \frac{\Gamma(x - t)}{\Gamma(x)} \leq 2x^{-t}$$

for all $x \geq m_0$. Thus,

$$\prod_{l=1}^{m} \frac{\Gamma(l - t)}{\Gamma(l)} \leq C(m_0, t) \cdot \frac{1}{((m_0 - 1)!)^t} \prod_{l=m_0}^{m} \frac{\Gamma(l - t)}{\Gamma(l)} \leq C(m_0, t) \frac{2^m}{(m!)^t}.$$

(2.22)
for \( m \geq m_0 \), where

\[
C(m_0, t) = ((m_0 - 1)!)^t \prod_{l=1}^{m_0-1} \left( 1 + \frac{\Gamma(l - t)}{\Gamma(l)} \right)
\]

is a constant depending on \( m_0 \) and \( t \) only. Therefore, from (2.22),

\[
\prod_{l=1}^{m} \frac{\Gamma(l - t)}{\Gamma(l)} \leq C(m_0, t) \frac{2^m}{(m!)^t} \tag{2.23}
\]

for all \( m \geq m_0 \). It is easy to check that (2.23) also holds when \( m \leq m_0 - 1 \). Thus, (2.23) is true for all \( m \).

On the other hand, from the assumption, we see that \( n - m \to \infty \) as \( n \to \infty \). By the same argument as in the above,

\[
\prod_{l=1}^{m} \frac{\Gamma(n - m + l)}{\Gamma(n - m + l - t)} \leq 2^m \prod_{l=1}^{m} (n - m + l)!^t = 2^m \prod_{j=n-m+1}^{n} j!^t = 2^m \left( \frac{n!}{(n-m)!} \right)^2 \tag{2.24}
\]

as \( n \) is sufficiently large. Combing (2.21), (2.23) and (2.24), we arrive at

\[
P\left( \frac{1}{n} \left| \log |\det (U_{[n,m]})| \right| > \epsilon \right) \leq C(m_0, t) e^{-2ntm} 4^m \left( \frac{n}{m} \right)^t
\]

as \( n \) is sufficiently large. By the Stirling formula \( m! = \sqrt{2\pi m} m^m e^{-m+\theta_m} \) with \( \theta_m \in (0, 1) \), we have that \( \left( \frac{n}{m} \right)^m \leq \left( \frac{n}{m} \right)^m \) for all \( 1 \leq m \leq n \). Consequently,

\[
P\left( \frac{1}{n} \left| \log |\det (U_{[n,m]})| \right| > \epsilon \right)
\leq C(m_0, t) \cdot \exp \left\{ - \left[ (2\epsilon) \frac{n}{m} - \log \frac{n}{m} - \log(4^{1/t}) \right] mt \right\} \leq C(m_0, t) \cdot e^{-3nt/2}
\]

as \( n \) is sufficiently large, where we use the inequality \( (2\epsilon) \frac{n}{m} - \log \frac{n}{m} - \log(4^{1/t}) \geq (1.5\epsilon) \frac{n}{m} \) in the last step since \( \lim_{n \to \infty} \frac{n}{m} = \infty \). Taking \( t = 3/4 \), we get (2.17). The proof is completed. \( \blacksquare \)

**Lemma 2.8** Let \( n \geq 2 \) and \( A \) be an \( n \times n \) unitary matrix. Write

\[
A = \left( \begin{array}{cc} A_1 & A_2 \\ A_3 & A_4 \end{array} \right)
\]

where \( A_1 \) and \( A_4 \) are square matrices. Then \( |\det(A_1)| = |\det(A_4)| \).
Proof. Let $A_1$ be $p \times p$ and $A_4$ be $q \times q$ with $p \geq 1$, $q \geq 1$ and $p+q = n$. Without loss of generality, assume $p \geq q$. Looking at the $(1,1)$-block of $A^*A = I_n$, we get $A_1^*A_1 + A_3^*A_3 = I_p$. It follows that

$$|\det(A_1)|^2 = \det(A_1^*A_1) = \det(I_p - A_3^*A_3) = \det(I_q - A_3^*A_3^*)$$  (2.25)

since the eigenvalues of $A_3^*A_3$ are the same as those of $A_3A_3^*$ plus 0 with $p-q$ fold. Looking at the $(2,2)$-block of $AA^* = I_n$, we have $A_3A_3^* + A_4A_4^* = I_q$. Thus

$$|\det(A_4)|^2 = \det(I_q - A_3A_3^*),$$

which together with (4.73) yields the conclusion. ■

Remark 2.1 Let $n \geq 2$ and $U_n = (u_{ij})$ be an $n \times n$ Haar unitary matrix and $U_{[n,n-1]}$ be the $(n-1) \times (n-1)$ upper-left corner of $U_n$. According to Lemma 2.8 and the distribution of the entries of $U_n$ (see, for example, Jiang 2010), we have

$$|\det(U_{[n,n-1]})|^2 = |u_{nn}|^2 = \frac{\xi_1^2 + \xi_2^2}{\sum_{i=1}^{2n} \xi_i^2}$$  (2.26)

where $\{\xi_1, \ldots, \xi_{2n}\}$ are independent and identically distributed random variables with distribution $N(0,1)$, and "\(\sim\)" means the last two random variables in (2.26) have the same probability distribution. Thus, $|\det(U_{[n,n-1]})|^2$ is asymptotically $\chi^2(2)/(2n)$ as $n$ is large. Heuristically, $|\det(U_{[n,n-1]})|$ is close to $|\det(U_n)| = 1$. However, (2.26) shows that $\det(U_{[n,n-1]})$ is in fact very small. This seems counterintuitive.

Proof of Theorem 2.1. Set $g(z) = \det(zI_m - U_{[n,m]}) = z^m + b_{m-1}z^{m-1} + \cdots + b_0$. By Lemma 4.69, to prove the theorem, it is enough to show that

$$\frac{1}{m} \log \left( \frac{1}{|b_0|} \sum_{j=0}^{m} |b_j| \right) \to 0 \ a.s.$$  (2.27)

as $n \to \infty$, where $b_m := 1$. 


First, for any $\epsilon > 0$,

$$P\left(\frac{1}{m^2} \sum_{j=0}^{m} |b_j| \geq \epsilon\right) \leq \frac{1}{m^2 \epsilon^2} \cdot E\left(\frac{1}{m} \sum_{j=0}^{m} |b_j|\right)^2 \leq \frac{1}{m^2 \epsilon^2} \cdot \frac{1}{m} \sum_{j=0}^{m} E(|b_j|^2)$$

by a convex inequality. From Lemma 2.6 we see that

$$E(|b_{m-j}|^2) = \left(\frac{m}{j}\right) \leq 1$$

for all $1 \leq j \leq m$. Recalling $b_m := 1$, the above two assertions give that

$$\sum_{n=1}^{\infty} P\left(\frac{1}{m_n^2} \sum_{j=0}^{m_n} |b_j| \geq \epsilon\right) \leq \sum_{n=1}^{\infty} \frac{1}{m_n^2 \epsilon^2} < \infty$$

for any $\epsilon > 0$ since $m_n/n \to 1$ as $n \to \infty$. By the Borel-Cantelli lemma, we have that

$$\frac{1}{m^2} \sum_{j=0}^{m} |b_j| = \frac{1}{m_n^2} \sum_{j=0}^{m_n} |b_j| \to 0 \text{ a.s.}$$

as $n \to \infty$. Since $b_m = 1$, it is not hard to see

$$\frac{1}{m} \log \sum_{j=0}^{m} |b_j| \to 0 \text{ a.s.} \quad (2.28)$$

as $n \to \infty$.

Second, notice $|b_0| = |\det(U_{[n,m]})|$. Write

$$U_n = \begin{pmatrix} U_{[n,m]} & * \\ * & W_{[n,m'_n]} \end{pmatrix}$$

where $W_{[n,m'_n]}$ is the $m'_n \times m'_n$ lower-right corner of $U_n$ with $m'_n := n - m$. By Lemma 2.8, $|b_0| = |\det(W_{[n,m'_n]})|$. From Lemma 2.7,

$$\frac{1}{n} \log |b_0| = \frac{1}{n} \log |\det(W_{[n,m'_n]})| \to 0 \text{ a.s.} \quad (2.29)$$

as $n \to \infty$ since $m'_n/n = 1 - (m/n) \to 0$ by the assumption. Then (2.28) and (2.29) yield (2.27). The proof is completed.  

\[\blacksquare\]
2.3 The Circular Law and Its Proof

To state the result for $\alpha = 0$, we need to review some terminology. Denote by $\mathcal{M}(\mathbb{C})$ the collection of the Borel probability measures defined on the complex plane $\mathbb{C}$. Let $\mu$ and $\nu$ be two probability measures on $\mathbb{C}$ (or $\mathbb{R}^2$). Define

$$\rho(\mu, \nu) = \sup_{\|f\|_L \leq 1} \left| \int_{\mathbb{C}} f(x) \mu(dx) - \int_{\mathbb{C}} f(x) \nu(dx) \right|,$$

where $f$ is a bounded Lipschitz function defined on $\mathbb{C}$ with $\|f\| = \sup_{x \in \mathbb{C}} |f(x)|$, and $\|f\|_L = \|f\| + \sup_{x \neq y} |f(x) - f(y)|/|x - y|$. This metric generates the topology of the weak convergence of the probability measures on $\mathbb{C}$ (see, e.g., chapter 11 from Dudley (2002)), that is, $\mu_n$ converges to $\mu$ weakly as $n \to \infty$ if and only if $\lim_{n \to \infty} \int f(x) \mu_n(dx) = \int f(x) \mu(dx)$ for every bounded and continuous function $f(x)$ defined on $\mathbb{C}$, and if and only if $\rho(\mu_n, \mu) \to 0$ as $n \to \infty$. We endow $\mathcal{M}(\mathbb{C})$ with the standard weak convergence topology, which is, as explained, the same one generated by $\rho(\mu, \nu)$ in (2.30).

**Theorem 2.2 (Circular Law)** Suppose $m \to \infty$ and $m/n \to 0$ as $n \to \infty$. Let $\mu_0$ be the uniform probability measure on $\{z \in \mathbb{C}; |z| \leq 1\}$ and

$$\mu_m := \frac{1}{m} \sum_{i=1}^{m} \delta_{\sqrt{\frac{m}{n}} \lambda_i}, \quad n \geq 1.$$
Then,
(i) $\rho(\mu_m, \mu_0) \to 0$ in probability as $n \to \infty$;
(ii) With probability one, $\mu_m$ converges weakly to $\mu_0$ as $n \to \infty$ if $\lim_{n \to \infty} m_n/\sqrt{\log n} = \infty$.

In view of the second part of the observation in (2.1), thinking that $m = 2$ “is of a small portion” of $n = 5$, we indeed see from Theorem 2.2 that “the eigenvalues are clustered closer to the origin.” Further, the theorem tells us that by magnifying the eigenvalues with $\sqrt{\frac{m}{n}}$ times, they are asymptotically distributed on the unit disc uniformly. See Figure 2.

Now we make some comments about Theorem 2.2.

First, by measuring the variation distance between the entries of $U_{[n,m]}$ and $m^2$ independent and identically distributed complex Gaussian random variables, Jiang (2009) showed that (i) of Theorem 2.2 holds when $m \to \infty$ and $m = o(\sqrt{n})$ as $n \to \infty$. Theorem 2.2 improves the order to $m = o(n)$. In fact, the order $m = o(n)$ is the best one to make the circular law hold since a different limit law in (4.50) appears when $m/n \to \alpha \in (0,1)$.

Second, let $Y_n = (y_{ij})$ be an $n \times n$ matrix where $y_{ij}$’s are independent and identically distributed random variables. Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of $Y_n$. The circular law problem, that is, the empirical distribution of $\lambda_1/\sqrt{n}, \cdots, \lambda_n/\sqrt{n}$ goes to the uniform distribution on $\{z \in \mathbb{C}; |z| \leq 1\}$ as $n \to \infty$, has been studied by some authors including, for example, Girko (1984a, 1984b), Bai (1997), Pan and Zhou (2010), Tao and Vu (2008, 2009, 2010). Theorem 2.2 presents a circular law when the entries of the matrix, unlike those of $Y_n$, are dependent random variables.

2.3.1 The Proof of The Circular Law

The main focus of the proof of Theorem 3.4 is to estimate $P(\rho(\mu_m, \mu_0) \geq \epsilon)$. By using an inequality we first transform the original problem to that on the Ginibre
ensemble. An important ingredient of the estimate is a large deviation for the

Suppose $X_n = (X_{ij})$ is an $n \times n$ matrix whose $n^2$ entries are independent and
identically distributed complex normal random variables with $E(X_{11}) = 0$ and
$E(|X_{11}|^2) = 1/n$. In other words, $X_{11}$ and $(\xi + i\eta)/\sqrt{2n}$ have the same distribution
where $\xi$ and $\eta$ are i.i.d. $N(0,1)$-distributed real random variables. In literature,
$X_n$ is called a complex Ginibre ensemble, see Ginibre (1965). Let $\lambda_1, \ldots, \lambda_n$ be
the (complex) eigenvalues of $X_n$. Their joint probability density function is given
by
\[
f(\lambda_1, \cdots, \lambda_n) = C_n \cdot e^{-n \sum_{i=1}^n |\lambda_i|^2} \cdot \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^2 d\lambda_1 \cdots d\lambda_n \tag{2.31}
\]
for all $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$, where $d\lambda_1 \cdots d\lambda_n := \prod_{i=1}^n d\text{Re}(\lambda_i) \prod_{i=1}^n d\text{Im}(\lambda_i)$ is the
Lebesgue integral in $\mathbb{R}^{2n}$ and
\[
C_n = \frac{n^{n(n+1)/2}}{\pi^n \prod_{k=1}^n k!}. \tag{2.32}
\]
See, for example, (1.35) and (1.36) from Ginibre (1965) (note the notation $dz\,dz^* =
2\,dx\,dy$ for $z = x + iy$ in the paper).

Let $\mu_0$ be the uniform probability distribution on the complex unit disc \( \{ z \in \mathbb{C}; |z| \leq 1 \} \). Given $\lambda_1, \cdots, \lambda_n$, define
\[
P_n = \frac{1}{n^2} \sum_{i=1}^n \delta_{\lambda_i} \tag{2.33}
\]
as the empirical distribution of $\lambda_1, \cdots, \lambda_n$. Recall $\mathcal{M}(\mathbb{C})$ above (2.30) and the
associated weak topology. We need the following large deviation result.

**Lemma 2.9** (*Theorem 9 from Hiai and Petz (1998))*: Suppose $\lambda_1, \cdots, \lambda_n$ have
the joint density function $f(\lambda_1, \cdots, \lambda_n)$ as in (2.31). Then \( \{ P_n; n \geq 1 \} \) satisfies
the large deviation principle with speed \{ $n^2$; $n \geq 1$ \} and a good rate function $I(\mu)$,
which takes the unique minimum at $\mu_0$ and $I(\mu_0) = 0$. In particular,
\[
\limsup_{n \to \infty} \frac{1}{n^2} \log P(P_n \in F) \leq -\inf_{\mu \in F} I(\mu) < 0
\]
for any closed set $F \subset \mathcal{M}(\mathbb{C})$ with $\mu_0 \notin F$. 
Let \( 1 \leq m < n \). Recall \( U_{[n,m]} \) is the \( m \times m \) upper-left corner of an \( n \times n \) Haar unitary matrix \( U_n \). It is known from Żyyczkowski and Sommers (2000) and Petz and Réffy (2005) that the joint probability density function of \( \lambda_1, \cdots, \lambda_m \), the eigenvalues of \( U_{[n,m]} \), is given by

\[
g(\lambda_1, \cdots, \lambda_m) = C_{[n,m]} \prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j|^2 \cdot \prod_{i=1}^{m} (1 - |\lambda_i|^2)^{n-m-1} d\lambda_1 \cdots d\lambda_m \tag{2.34}
\]

with

\[
\frac{1}{C_{[n,m]}} = \pi^m m! \prod_{j=0}^{m-1} \left( n - m + j - 1 \right)^{-1} \frac{1}{n - m + j}.
\]

\[
= \pi^m \prod_{j=1}^{m} j! \prod_{j=n-m}^{n-1} \left[ (n - m - 1)! \right]^{-1}. \tag{2.35}
\]

**Lemma 2.10** Let \( \{a_{n,i}; 1 \leq i \leq n, n \geq 1\} \) and \( \{t_n; n \geq 1\} \) be complex numbers such that \( \lim_{n \to \infty} t_n = 1 \). Set

\[
\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{a_{n,i}} \quad \text{and} \quad \mu'_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{(t_na_{n,i})}
\]

for \( n \geq 1 \). If \( \mu_n \) converges weakly to a Borel probability measure \( \nu \) on \( \mathbb{C} \), then \( \mu'_n \) also converges weakly to \( \nu \).

**Proof.** Let \( f(x) \) be a real-valued function defined on \( \mathbb{C} \) with

\[
\sup_{x \in \mathbb{C}} |f(x)| = 1 \quad \text{and} \quad |f(x) - f(y)| \leq \alpha|x - y| \tag{2.36}
\]

for all \( x, y \in \mathbb{C} \), where \( \alpha \in (0, \infty) \) is a constant. We need to show

\[
\frac{1}{n} \sum_{i=1}^{n} f(t_na_{n,i}) \to \int_{\mathbb{C}} f(x) \nu(dx)
\]

as \( n \to \infty \). By the given condition, the above is true with \( t_n \equiv 1 \). So it suffices to show

\[
H_n := \frac{1}{n} \sum_{i=1}^{n} |f(t_na_{n,i}) - f(a_{n,i})| \to 0 \tag{2.37}
\]
as \( n \to \infty \). For any \( r > 0 \), set \( F_r = \{ x \in \mathbb{C}; |x| \geq r \} \). Using (2.36) we have

\[
\frac{1}{n} \sum_{i=1}^{n} |f(t_n a_{n,i}) - f(a_{n,i})| \leq \frac{2}{n} \sum_{i=1}^{n} I\{|a_{n,i}| \geq r\} + \alpha r|t_n - 1| = 2 \mu_n(F_r) + \alpha r|t_n - 1|.
\]

Then \( \limsup_{n \to \infty} H_n \leq 2 \nu(F_r) \). We get (2.37) by letting \( r \to +\infty \).

**Proof of Theorem 2.2.** Set \( t_n = \sqrt{(n - m - 1)/n} \) and

\[
\mu'_m = \frac{1}{m} \sum_{i=1}^{m} \delta_{(t_n \sqrt{n/m \lambda_i})}
\]

for \( n \geq 1 \). We will first show that, for any \( \epsilon > 0 \), there exists a constant \( \tau = \tau(\epsilon) > 0 \) such that

\[
P(\rho(\mu'_m, \mu_0) \geq \epsilon) \leq e^{-\tau m^2}
\]

as \( n \) is sufficiently large, where \( \rho \) is as in (2.30).

(i) If (2.39) is true, then \( \rho(\mu'_m, \mu_0) \to 0 \) in probability as \( n \to \infty \). Recall that random variable \( T_n \to T \) in probability as \( n \to \infty \) if and only if for any subsequence \( \{n_k\} \) of \( \{n\} \), there is a further subsequence \( \{n_{k_j}\} \) such that \( T_{n_{k_j}} \to T \) almost surely as \( j \to \infty \). Notice \( \lim_{n \to \infty} t_n = 1 \) since \( m = m_n = o(n) \) as \( n \to \infty \). By Lemma 2.10, we have that \( \rho(\mu_m, \mu_0) \to 0 \) in probability as \( n \to \infty \).

(ii) If (2.39) is true, by the assumption \( \lim_{n \to \infty} m_n/\sqrt{\log n} = \infty \), we see that

\[
\sum_{n \geq m_1} P(\rho(\mu'_m, \mu_0) \geq \epsilon) < \infty
\]

for any \( \epsilon > 0 \). It follows from the Borel-Cantelli lemma that, with probability one, \( \rho(\mu'_m, \mu_0) \to 0 \) as \( n \to \infty \), equivalently, \( \mu'_m \) converges weakly to \( \mu_0 \) as \( n \to \infty \). By Lemma 2.10, with probability one, \( \mu_m \) converges weakly to \( \mu_0 \) as \( n \to \infty \).
Now we prove (2.39). From (2.34), we have

\[
P(\rho(\mu'_m, \mu_0) \geq \epsilon) = C_{[n,m]} \int_{\rho(\rho'_m, \mu_0) \geq \epsilon} \prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j|^2 \cdot \prod_{i=1}^{m} (1 - |\lambda_i|^2)^{n-1} \prod_{i=1}^{m} d\lambda_i \leq C_{[n,m]} \int_{\rho(\rho'_m, \mu_0) \geq \epsilon} \prod_{1 \leq i < j \leq m} |\lambda_i - \lambda_j|^2 \cdot e^{-(n-m-1) \sum_{i=1}^{m} |\lambda_i|^2} \prod_{i=1}^{m} d\lambda_i
\]

where \( \prod_{i=1}^{m} d\lambda_i = \prod_{i=1}^{m} d\text{Re}(\lambda_i) \prod_{i=1}^{m} d\text{Im}(\lambda_i) \), and \( C_{[n,m]} \) is as in (2.35) and the inequality \( 1 - x \leq e^{-x} \) for any \( x \in \mathbb{R} \) is used in the last step. Set \( y_i = s_n \lambda_i \) for \( i = 1, 2, \ldots, m \), where \( s_n = \sqrt{(n-m-1)/m} \). Then

\[
P(\rho(\mu'_m, \mu_0) \geq \epsilon) \leq \frac{(s_n)^{-m(m+1)} C_{[n,m]}}{C_m} \int_{\rho(\rho'_m, \mu_0) \geq \epsilon} \cdot \prod_{1 \leq i < j \leq m} |y_i - y_j|^2 \cdot e^{-m \sum_{i=1}^{m} |y_i|^2} \prod_{i=1}^{m} d\lambda_i \tag{2.40}
\]

where \( C_m \) is as in (2.32) and

\[
P_m := \frac{1}{m} \sum_{i=1}^{m} \delta_{y_i}
\]

is the same as the notation in (2.33). Observe that the integral in (2.40) is equal to \( P(P_m \in F) \) where the underlying probability measure \( P \) has density function \( f(\lambda_1, \ldots, \lambda_m) \) as in (2.31) and (2.32), and where \( F := \{ \mu \in \mathcal{M}(\mathbb{C}) ; \rho(\mu, \mu_0) \geq \epsilon \} \) is a closed set in the weak convergence topology and \( \mu_0 \notin F \). By Lemma 2.9, there exists a constant \( \tau = \tau(\epsilon) > 0 \) such that

\[
P(P_m \in F) \leq e^{-2m^2 \tau}
\]

as \( n \) is sufficiently large, and hence \( m = m_n \) is sufficiently large. This combining with (2.40) gives

\[
P(\rho(\mu'_m, \mu_0) \geq \epsilon) \leq \frac{(s_n)^{-m(m+1)} C_{[n,m]}}{C_m} e^{-2m^2 \tau} \tag{2.41}
\]
as \( n \) is sufficiently large. Note that

\[
\frac{(s_n)^{-m(m+1)}C_{[n,m]}}{C_m} = (s_n)^{-m(m+1)} \cdot \frac{\prod_{j=n-m}^{n-1} j!}{(n-m-1)!} \cdot \frac{1}{m^{m(m+1)/2}}. \tag{2.42}
\]

Observe that \( j! = (n-m-1)!(n-m) \cdots j \leq (n-m-1)!n^{j-n+m+1} \) for all \( n-m \leq j \leq n-1 \). Recalling \( s_n = \sqrt{(n-m-1)/m} \), we get from (3.3) that

\[
\frac{(s_n)^{-m(m+1)}C_{[n,m]}}{C_m} \leq (s_n)^{-m(m+1)} \cdot \frac{n^{m(m+1)/2}}{m^{m(m+1)/2}} = \left(1 - \frac{m+1}{n}\right)^{-m(m+1)/2}
\]

which joint with (2.41) yields

\[
\limsup_{n \to \infty} \frac{1}{m^2} \log P(\rho'(\mu', \mu_0) \geq \epsilon)
\]

\[
\leq \limsup_{n \to \infty} \left(-2\tau - \frac{m+1}{2m} \log \left(1 - \frac{m+1}{n}\right)\right) = -2\tau
\]

since \( m = m_n = o(n) \). Consequently,

\[
P(\rho'(\mu', \mu_0) \geq \epsilon) \leq e^{-\tau m_n^2}
\]

as \( n \) is sufficiently large. Thus (2.39) is obtained. The proof is completed. \(\blacksquare\)
Chapter 3

Approximation of Rectangular Beta-Laguerre Ensembles and Large Deviations

3.1 Introduction

With the development of modern technology, high-dimensional datasets appear very frequently in different scientific disciplines such as climate studies, financial data, information retrieval/search engines and functional data analysis. The corresponding statistical problems have the feature that the dimension $p$ is possibly larger than the sample size $n$. In such cases the classical statistical procedures based on fixed $p$ and large $n$ are no more applicable. The applications thus request new theories. See, for example, Candes and Tao (2005), Donoho et al (2006), Cai and Jiang (2011, 2012), and Vershynin (2012) for progress in this area.

In this chapter we study the spectral properties of a Wishart matrix formed by a random sample of $p$-dimensional data with sample size $n$, where $p$ is larger than $n$ as a special case of beta-Laguerre Ensembles. Wishart matrices are very popular and useful objects in multivariate analysis, see, for example, the classical
books by Muirhead (1982) and Anderson (1984). It usually comes from the following formulation in statistics. Let \( y_1, \ldots, y_m \) be i.i.d. random variables with the \( p \)-dimensional multivariate normal distribution \( N_p(\mu, I_p) \). Recall the sample covariance matrix

\[
S = \frac{1}{m} \sum_{i=1}^{m} (y_i - \bar{y})(y_i - \bar{y})^* \quad \text{where} \quad \bar{y} = \frac{1}{m} \sum_{i=1}^{m} y_i.
\]  

(3.1)

Then \( mS \) and \( W := XX^* \) have the same distribution, where \( X = (x_{ij})_{n \times p} \), the random variables \( x_{ij} \)'s are i.i.d. with distribution \( N(0, 1) \) and \( n = m - 1 \). The matrix \( W \) is referred to as the real Wishart matrix (\( \beta = 1 \)). If \( x_{ij} \) are i.i.d. with the standard complex or quaternion normal distribution, then \( W \) is a complex Wishart matrix (\( \beta = 2 \)) or quaternion Wishart matrix (\( \beta = 4 \)).

Assume \( p > n \). Let \( \lambda_1 > \cdots > \lambda_n > 0 \) be the positive eigenvalues of \( W \), which are the same as the \( n \) eigenvalues of \( XX^* \). Write \( \lambda = (\lambda_1, \cdots, \lambda_n) \). It is known that the density function of \( \lambda \) is given by

\[
f_{n,\beta}(\lambda) = c_{n}^{\beta,p} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^{\beta} \cdot \prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(p-n+1)} \cdot e^{-\frac{1}{2} \sum_{i=1}^{n} \lambda_i}
\]  

(3.2)

for all \( \lambda_1 > 0, \cdots, \lambda_n > 0 \), where

\[
c_{n}^{\beta,p} = 2^{-\frac{\beta}{2}np} \prod_{j=1}^{n} \frac{\Gamma(1 + \frac{\beta}{2})}{\Gamma(1 + \frac{\beta}{2}j)\Gamma(\frac{\beta}{2}(p-n+j))}.
\]  

(3.3)

See, for example, James (1964) and Muirhead (1982) for the cases \( \beta = 1 \) and \( 2 \), and Macdonald (1995) and Edelman and Rao (2005) for \( \beta = 4 \). The function \( f_{n,\beta}(\lambda) \) in (3.2), being a probability density function for any \( \beta > 0 \), is called the \( \beta \)-Laguerre ensemble in literature. See, for example, Dumitriu (2003) and Dumitriu and Edelman (2006).

We study the properties of \( \lambda = (\lambda_1, \cdots, \lambda_n) \) for all \( \beta > 0 \). Precisely, there are two objectives. First, we show in Theorem 3.1 that, when \( p \) is much larger than \( n \) in a certain scale, a “normalized” \( \beta \)-Laguerre ensemble can be roughly thought as a \( \beta \)-Hermite ensemble with density function

\[
f_{\beta}(\lambda) = K_{n}^{\beta} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^{\beta} \cdot e^{-\frac{1}{2} \sum_{i=1}^{n} \lambda_i^2}
\]  

(3.4)
for all $\lambda = (\lambda_1, \cdots, \lambda_n) \in \mathbb{R}^n$, where

$$K_{n}^{\beta} = (2\pi)^{-n/2} \prod_{j=1}^{n} \frac{\Gamma(1 + \frac{\beta}{2})}{\Gamma(1 + \frac{\beta}{2}j)}.$$  \hspace{2cm} (3.5)

The eigenvalues $\lambda_1, \cdots, \lambda_n$ of the Gaussian orthogonal ensemble (GOE), Gaussian unitary ensemble (GUE) and Gaussian symplectic ensemble (GSE) have the joint density function $f_{\beta}(\lambda)$ as in (3.4) with $\beta = 1, 2$ and $4$, respectively. See, e.g., chapter 17 from Mehta (1991) for more details.

By using Theorem 3.1 mentioned above (3.4) and some known results on $\beta$-Hermite ensembles, under the assumption that $p$ is much larger than $n$ in a certain scale, we obtain the following new results:

(i) The largest eigenvalue $\lambda_{\max}$ and smallest eigenvalue $\lambda_{\min}$ in the $\beta$-Laguerre ensemble as $\beta = 2$ are asymptotically independent (Proposition 3.1).

(ii) The condition number $\kappa_n$ (see the definition in (3.14)) of an $n \times p$ matrix $(x_{ij})$ ($x_{ij}$’s are i.i.d. centered complex Gaussian random variables), when suitably normalized, converges weakly to $U + V$ where $U$ and $V$ are independent random variables with a common Tracy-Widom law (Corollary 3.1). This is much different from the exact square case that $n = p$ studied by Edelman (1988): $\kappa_n/n$ converges weakly to a distribution with density function $h(x) = 8x^{-3}e^{-4/x^2}$ for $x > 0$. Based on this result, a spherical test in statistics is proposed below Corollary 3.1.

(iii) A linear transform of the smallest eigenvalue of the $\beta$-Laguerre ensemble converges to the $\beta$-Tracy-Widom law for any $\beta > 0$ (Proposition 3.2). The counterpart for the largest eigenvalues was studied by Ramírez et al (2011).

It is worthwhile to mention that the condition number $\kappa_n$ mentioned in (ii) is an important quantity in the field of numerical analysis dated back to Von Neumann and Goldstine (1963).

In the second part of this chapter we study the large deviations for the eigenvalues of the $\beta$-Laguerre ensembles when $p$ is much larger than $n$. The large deviation for the eigenvalues of random matrices is one of active research areas in random matrix theory. See, for example, a survey paper by Guionnet (2004) and some chapters from Hiai and Petz (2006) and Anderson et al (2009). In particular, Ben
Arous and Guionnet (1997) investigate the large deviation for Wigner’s semi-circle law; Ben Arous, Dembo and Guionnet (2001) and Anderson et al (2009) study the largest eigenvalues of Wigner and Wishart matrices; A corollary from Hiai and Petz (1998) says that a normalized empirical distribution \( \mu_n \) of the positive eigenvalues of real Wishart matrix \( X_{n \times p}X_{n \times p}^* \) follows the large deviation principle (LDP) such that

\[
\limsup_{n \to \infty} \frac{1}{p^2} \log P(\mu_n \in F) \leq -\inf_{\nu \in F} I(\nu) \quad \text{and} \quad (3.6)
\]

\[
\liminf_{n \to \infty} \frac{1}{p^2} \log P(\mu_n \in G) \geq -\inf_{\nu \in G} I(\nu) \quad (3.7)
\]

for every closed set \( F \) and open set \( G \) under the topology of weak convergence of probability measures on \( \mathbb{R} \), where \( \mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i/p} \), the eigenvalues \( \lambda_1, \ldots, \lambda_n \) have joint density \( f_{n,1}(\lambda) \) as in (3.2) with \( n/p \to \gamma \in (0,1] \), and

\[
I(\nu) = -\frac{\gamma^2}{2} \iint \log |x - y| d\nu(x)d\nu(y) + \frac{\gamma}{2} \int (x - (1 - \gamma) \log x) d\nu(x) + \text{Const}
\]

which takes the minimum value zero at the Marchenko-Pastur law with density function

\[
h(x) = \frac{1}{2\pi \gamma x} \sqrt{(x - \gamma_1)(\gamma_2 - x)} \quad (3.8)
\]

for \( x \in [\gamma_1, \gamma_2] \) and \( \gamma_1 = (\sqrt{\gamma} - 1)^2 \) and \( \gamma_2 = (\sqrt{\gamma} + 1)^2 \). For the general framework of the large deviation principle, its connection to the subjects of mathematics, physics, statistics and engineering, see, for example, Shwartz and Weiss (1995), Ellis (2011), and Dembo and Zeitouni (2009).

When \( p/n \to \infty \), the LDP problem for \( \mu_n \) in (3.6) and (3.7) has been open until now. In fact, we resolve the problem in Theorem 3.4 under the assumption that both \( p \) and \( n \) are large with \( p/n^2 \to \infty \). Contrary to the Marchenko-Pastur law stated in (3.8), we show that the rate function in Theorem 3.4 takes the minimum value at the semi-circle law.

The large deviation principles for the largest eigenvalue \( \lambda_{\text{max}} \) and the smallest eigenvalue \( \lambda_{\text{min}} \) of the \( \beta \)-Laguerre ensemble are also studied in Theorems 3.2 and 3.3 as \( p/n \to \infty \). Their rate functions are explicit.
The rest of this chapter is organized as follows. In Section 3.2 we give a theorem that the \( \beta \)-Laguerre ensemble converges to the \( \beta \)-Hermite ensemble as \( p \) is much larger than \( n \) and present some implications; in section 3.3 we give three theorems about the large deviations for the largest eigenvalues, the smallest eigenvalues and the empirical distributions of the eigenvalues of the \( \beta \)-Laguerre ensemble as \( p \) is much larger than \( n \) in a certain scale. The proofs of the results stated in Sections 3.2 and 3.3 are given in Sections 3.4 and 3.5, respectively.

The reader is warned that the notation \( \mu \) or \( \mu_n \) throughout this chapter sometimes represents a probability measure, a mean value or an eigenvalue in different occasions, but this will not cause confusions from the context.

### 3.2 Convergence of Laguerre Ensembles to Hermite Ensembles and Its Applications

Let \( \mu \) and \( \nu \) be probability measures on \((\mathbb{R}^k, \mathcal{B})\), where \( k \geq 1 \) and \( \mathcal{B} \) is the Borel \( \sigma \)-algebra on \( \mathbb{R}^k \). The variation distance \( \|\mu - \nu\| \) is defined by

\[
\|\mu - \nu\| = 2 \cdot \sup_{A \in \mathcal{B}} |\mu(A) - \nu(A)| = \int_{\mathbb{R}^k} |f(x) - g(x)| \, dx_1 \cdots dx_k \tag{3.9}
\]

if \( \mu \) and \( \nu \) have density functions \( f(x) \) and \( g(x) \) with respect to the Lebesgue measure. For a random vector \( Z \), we use \( L(Z) \) to denote its probability distribution. The notation \( a_n \gg b_n \) means \( \lim_{n \to \infty} a_n/b_n = +\infty \).

**Theorem 3.1** Let \( \lambda = (\lambda_1, \cdots, \lambda_n) \) be random variables with density function \( f_{n,\beta}(\lambda) \) as in (3.2) and \( \mu = (\mu_1, \cdots, \mu_n) \) be random variables with density \( f_\beta(\mu) \) as in (3.4). Set \( x_i = \sqrt{\frac{p}{2\beta}}(\frac{\lambda_i}{p} - \beta) \) for \( 1 \leq i \leq n \). Then \( \|L((x_1, \cdots, x_n)) - L((\mu_1, \cdots, \mu_n))\| \to 0 \) if (i) \( n \to \infty \) and \( p = p_n \gg n^3 \) or (ii) \( n \) is fixed and \( p \to \infty \).

Roughly speaking, Theorem 3.1 says that a “very rectangular-shaped” \( \beta \)-Laguerre ensemble is essentially a \( \beta \)-Hermite ensemble. Combining Theorem 3.1
with some known results on the $\beta$-Hermite ensembles, we obtain several new results. To state them, let us first review the Tracy-Widom distributions. Set
\[
F_2(x) = \exp \left( - \int_x^\infty (y - x)q^2(y)\,dy \right), \quad x \in \mathbb{R}, \tag{3.10}
\]
where $q$ is the unique solution to the Painlevé II differential equation
\[
q''(x) = xq(x) + 2q^3(x) \tag{3.11}
\]
satisfying the boundary condition $q(x) \sim Ai(x)$ as $x \to \infty$, where $Ai(x)$ is the Airy function. It is known from Hastings and McLeod (1980) that
\[
q(x) = \sqrt{-\frac{x}{2}} \left( 1 + \frac{1}{8x^3} + O\left(\frac{1}{x^6}\right) \right)
\]
as $x \to -\infty$. The distributions for the orthogonal and symplectic cases (Tracy and Widom (1996)) are
\[
F_1(x) = \exp \left( - \frac{1}{2} \int_x^\infty q(y)\,dy \right) \left( F_2(x) \right)^{1/2} \quad \text{and} \quad \tag{3.12}
\]
\[
F_4(x/\sqrt{2}) = \cosh \left( \frac{1}{2} \int_x^\infty q(y)\,dy \right) \left( F_2(x) \right)^{1/2} \tag{3.13}
\]
for all $x \in \mathbb{R}$, where $\cosh t = (e^t + e^{-t})/2$ for $t \in \mathbb{R}$.

Our first result following from Theorem 3.1 is on complex Wishart matrices ($\beta = 2$). In fact, data matrices with complex-valued entries arise frequently, for example, in signal processing applications (e.g., Tulino and Verdu (2004)) and statistics (e.g., James (1964) and Picinbono (1996)). Given $\lambda_1, \ldots, \lambda_n$, set $\lambda_{\min} = \min\{\lambda_1, \ldots, \lambda_n\}$ and $\lambda_{\max} = \max\{\lambda_1, \ldots, \lambda_n\}$.

**Proposition 3.1** Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be random variables with density $f_{n,\beta}(\lambda)$ as in (3.2) with $\beta = 2$. Set $\mu_{n,1} = 2p - 4\sqrt{np}$, $\mu_{n,2} = 2p + 4\sqrt{np}$ and $\sigma_n = 2\sqrt{p}n^{-1/6}$. If $p = p_n \to \infty$ and $p \gg n^3$, then $((\lambda_{\min} - \mu_{n,1})/\sigma_n, (\lambda_{\max} - \mu_{n,2})/\sigma_n) \in \mathbb{R}^2$ converge weakly to $(-U, V)$, where $U$ and $V$ are i.i.d. with distribution function $F_2(x)$ as in (3.10).
Basor et al (2012) heuristically show that Proposition 3.1 is true as $n/p \to c$. Our result above is rigorous. It remains open at this moment if Proposition 3.1 is still true for $\beta \neq 2$ under the assumption $p/n \to \gamma \in [1, \infty]$.

Let $X = (x_{ij})_{n \times p}$ and $x_{ij}$’s be i.i.d. complex random variables with the distribution of $(\xi + i\eta)/\sqrt{2}$ where $\xi$ and $\eta$ are i.i.d. with $N(0,1)$-distribution. Suppose $\lambda_{\text{max}}$ and $\lambda_{\text{min}}$ are the largest and smallest eigenvalues of the matrix $XX^*$ (the density of the eigenvalues of this matrix corresponds to $\beta = 2$ in (3.2)). Recall the condition number defined by

$$\kappa_n := \left(\frac{\lambda_{\text{max}}}{\lambda_{\text{min}}}\right)^{1/2}. \tag{3.14}$$

An immediate consequence of Proposition 3.1 is the following result about $\kappa_n$.

**Corollary 3.1** Let $\lambda = (\lambda_1, \cdots, \lambda_n)$ be the eigenvalues of $XX^*$ stated above. If $p \gg n^3$, then $\alpha_n (\kappa_n - \beta_n)$ converges weakly to $U + V$, where $\alpha_n = 2\sqrt{pn^{1/6}}$, $\beta_n = 1 + 2\sqrt{\frac{2}{p}}$, and $U$ and $V$ are i.i.d. with distribution function $F_2(x)$ as in (3.10).

In the exact square case that $p = n$, Edelman (1988) proves that $\kappa_n/n$ converges weakly to a distribution with density function $h(x) = 8x^{-3}e^{-4/x^2}$ for $x > 0$. In the rectangular case such that $p \gg n^3$, Corollary 3.1 shows a very different behavior of $\kappa_n$.

Let $Y = \xi + i\eta$ be a multivariate complex normal distribution where $\xi \sim N_p(0, \Sigma_1)$ and $\eta \sim N_p(0, \Sigma_2)$ are independent. Consider the spherical test $H_0 : \Sigma_1 = \Sigma_2 = \rho I_p$ vs $H_a : H_0$ is not true, where $\rho > 0$ is not specified. Let $Y_1, \cdots, Y_n$ be a random sample from the population distribution of $Y$ with $p > n$. The classical likelihood ratio test (see, e.g., Muirhead (1982)) does not work here simply because $p > n$. In this situation, if $p$ is much larger than $n$ with $p \gg n^3$, we can use Corollary 3.1 to do the test: set $X = X_{n \times p} = (Y_1, \cdots, Y_n)'$. Then the $n$ positive eigenvalues of $X^*X/\rho^2$ have the joint density function $f_{n,\beta}(\lambda)$ as in (3.2) with $\beta = 2$. Recall that $\lambda_{\text{max}}$ and $\lambda_{\text{min}}$ are the largest and smallest positive eigenvalues of $X^*X$, respectively. Then, $\kappa_n = (\lambda_{\text{max}}/\lambda_{\text{min}})^{1/2}$ does not depend
on the unknown parameter $\rho$. By Corollary 3.1, the region to reject $H_0$ with an asymptotic $1 - \alpha$ confidence level is \( \{ \alpha_n|\kappa_n - \beta_n| > s \} \), where $s > 0$ satisfies $P(|U + V| > s) = \alpha$. The value of $s$ can be calculated through a numerical method by using (3.10), (3.11) and the independence between $U$ and $V$.

Now we study the limiting distribution of $\lambda_{\min}$ in the $\beta$-Laguerre ensemble for all $\beta > 0$. To do so, consider the random operator

$$
\mathcal{H}_\beta = -\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} b'_x \tag{3.15}
$$

where $b_x$ is a standard Brownian motion on $[0, +\infty)$ ($b'_x$ is not the derivative of $b_x$ since it is not differentiable almost everywhere). We use equation (3.15) in the following sense. For $\lambda \in \mathbb{R}$ and function $\psi(x)$ defined on $[0, +\infty)$ with $\psi(0) = 0$ and $\int_0^\infty ((\psi')^2 + (1 + x)\psi^2) \, dx < \infty$, we say $(\psi, \lambda)$ is an eigenfunction/eigenvalue pair for $\mathcal{H}_\beta$ if

$$
\int_0^\infty \psi^2(x) \, dx = 1 \quad \text{and} \quad \psi''(x) = \frac{2}{\sqrt{\beta}} \psi(x)b'_x + (x - \lambda)\psi(x)
$$

holds in the sense of integration-by-parts, that is,

$$
\psi'(x) - \psi'(0) = \frac{2}{\sqrt{\beta}} \psi(x)b_x + \int_0^x -\frac{2}{\sqrt{\beta}} b_y\psi'(y) \, dy + \int_0^x (y - \lambda)\psi(y) \, dy. \tag{3.16}
$$

Theorem 1.1 from Ramírez et al (2011) shows that, with probability one, for each $k \geq 1$, the set of eigenvalues of $\mathcal{H}_\beta$ has well-defined $k$-lowest eigenvalues $(\Lambda_0, \ldots, \Lambda_{k-1})$. Our result on $\lambda_{\min}$ of a $\beta$-Laguerre ensemble is given next.

**Proposition 3.2** Let $\beta > 0$ and $\lambda = (\lambda_1, \ldots, \lambda_n)$ be random variables with density $f_{n,\beta}(\lambda)$ as in (3.2). Set $\mu_n = \beta(p - 2\sqrt{n p})$ and $\sigma_n = \beta\sqrt{pn^{-1/6}}$. If $p = p_n \to \infty$ and $p \gg n^3$, then, as $n \to \infty$, $(\lambda_{\min} - \mu_n)/\sigma_n$ converges weakly to the distribution of $\Lambda_0$.

It is known from Ramírez et al (2011) that $-\Lambda_0$ has the distribution $F_\beta(x)$ as in (4.79), (3.10) and (3.13) for $\beta = 1, 2$ and 4.
Under the less restrictive condition that \( p \gg n \), Paul (2011) obtains Proposition 3.2 for \( \beta = 1 \) and 2. Further, assuming \( p/n \to \gamma \in (1, \infty) \), Baker et al (1998) show that, if \( \beta = 2 \), then \( (\lambda_{\text{min}} - \nu_n)/\tau_n \) converges weakly to the distribution function \( 1 - F_2(-x) \) (the distribution of \(-\Lambda_0\) for \( \beta = 2 \)), where \( \nu_n \) and \( \tau_n \) are normalizing constants. Ma (2010) obtains a similar result for \( \beta = 1 \). Here, Proposition 3.2 holds for any \( \beta > 0 \).

For largest eigenvalue \( \lambda_{\text{max}} \), Johansson (2000), Johnstone (2001) and Karoui (2003) obtain its limiting distribution as \( \beta = 1, 2, 4 \) and \( \gamma \in [0, \infty] \). For general \( \beta > 0 \), the limiting distribution of \( \lambda_{\text{max}} \) is obtained by Ramírez et al (2011) for the \( \beta \)-Laguerre ensembles (that is, \( \lambda = (\lambda_1, \cdots, \lambda_n) \) has density \( f_{n,\beta}(\lambda) \) as in (3.2)) as \( p/n \to \gamma \in [1, \infty) \). We derive the asymptotic distribution of \( \lambda_{\text{min}} \) for the same \( \beta \)-Laguerre ensemble when \( p \gg n^3 \) in Proposition 3.2. At this point it is not known if a result similar to Proposition 3.2 still holds as \( p/n \to \gamma \in (0, \infty) \).

Although the proof of Theorem 3.1 suggests that the order of \( p \gg n^3 \) in Theorem 3.1 is the best one to make the approximation hold, the orders appearing in Propositions 3.1 and 3.2 and Corollary 3.1 could be relaxed. This is because Theorem 3.1 is a uniform approximation, and the three results are specific cases. One can see improvements in a different but similar situations in Dong et al (2012).

3.3 Large Deviations for Eigenvalues

In this section we study the large deviations for three basic statistics as \( p \gg n \) : the largest eigenvalue \( \lambda_{\text{max}} \), the smallest eigenvalue \( \lambda_{\text{min}} \) and the empirical distribution of \( \lambda_1, \cdots, \lambda_n \) which come from a \( \beta \)-Laguerre ensemble. One can check, for example, Dembo and Zeitouni (2009) for the definition of the large deviation principle (LDP). The first one is about the largest eigenvalue.

**Theorem 3.2** Suppose \( \lambda_1, \cdots, \lambda_n \) have the density \( f_{n,\beta}(\lambda) \) as in (3.2). Assume \( p = p_n \gg n \) as \( n \to \infty \). Then, \( \left\{ \frac{\lambda_{\text{max}}}{p}; n \geq 2 \right\} \) satisfies the LDP with speed
\{p_n; n \geq 2\} and rate function \(I(x)\) where
\[
I(x) = \begin{cases} 
\frac{x-\beta}{2} - \frac{\beta}{2} \log \frac{x}{\beta}, & \text{if } x \geq \beta; \\
+\infty, & \text{if } x < \beta.
\end{cases}
\]

For the smallest eigenvalue, we have the following.

**Theorem 3.3** Suppose \(\lambda_1, \ldots, \lambda_n\) have the density \(f_{n,\beta}(\lambda)\) as in (3.2). Assume 
\(p = p_n \gg n\) as \(n \to \infty\). Then, \(\{\frac{\lambda_{\min}}{p}; n \geq 2\}\) satisfies the LDP with speed 
\(\{p_n; n \geq 2\}\) and rate function \(I(x)\) where
\[
I(x) = \begin{cases} 
x - \frac{\beta}{2} - \frac{\beta}{2} \log \frac{x}{\beta}, & \text{if } 0 < x \leq \beta; \\
+\infty, & \text{otherwise.}
\end{cases}
\]

Since \(\lambda_{\max}/p\) and \(\lambda_{\min}/p\) are all positive, it can be seen from Theorems 3.2 and 3.3 that the two rate functions on \((0, \infty)\) look “symmetric” with respect to the line \(x = \beta\). When \(p/n \to (0, \infty)\), the large deviations for \(\lambda_{\max}/p\) is known, see, for example, Anderson, Guionnet and Zeitouni (2009).

Now we consider the large deviation for the empirical distribution of the eigenvalues.

**Theorem 3.4** Given \(\beta > 0\), let \(\lambda = (\lambda_1, \ldots, \lambda_n)\) have the joint density function \(f_{n,\beta}(\lambda)\) as in (3.2). Set \(x_i = \sqrt{\frac{p}{2\beta}} (\frac{\lambda_i}{p} - \beta)\) for \(1 \leq i \leq n\) and
\[
\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\frac{x_i}{\sqrt{n}}}, \tag{3.17}
\]

If \(p = p_n \to \infty\) and \(p \gg n^2\), then \(\{\mu_n; n \geq 2\}\) satisfies the LDP with speed \(\{n^2\}\) and rate function \(I_\beta(\nu)\), where
\[
I_\beta(\nu) = \frac{1}{2} \int_{\mathbb{R}^2} g(x, y) \nu(dx) \nu(dy) + \frac{\beta}{4} \log \frac{\beta}{2} - \frac{3}{8} \beta \tag{3.18}
\]

and
\[
g(x, y) = \begin{cases} 
\frac{1}{2}(x^2 + y^2) - \beta \log|x-y|, & \text{if } x \neq y; \\
+\infty, & \text{otherwise.}
\end{cases} \tag{3.19}
\]
As stated in Theorem 1.3 from Ben Arous and Guionnet (1997), $I_\beta(\nu)$ is the rate function of the large deviation for $\mu_n$ in (3.17) with $p = n$ when $x_1, \cdots, x_n$ come from a $\beta$-Hermite ensemble with density $f_\beta(x)$ as in (3.4), and the rate function $I_\beta(\nu)$ takes the unique minimum value 0 at the semi-circle law with density function $g_\beta(x) = (\beta \pi)^{-1} \sqrt{2\beta - x^2}$ for any $|x| \leq \sqrt{2\beta}$ and $\beta > 0$. This fact implies the law of large numbers: under the setting in Theorem 3.4, with probability one, $\mu_n$ converges weakly to a probability distribution with density $g_\beta(x)$. When $\beta = 1$, that is, the underlying matrix is the real Wishart matrix, Bai and Yin (1988) show the law of large numbers with the relaxed condition $p \gg n$. When $\beta \neq 1$, the law of large numbers is new.

Theorem 3.4 is consistent with Theorem 3.1 which says that the $\beta$-Laguerre ensemble is “essentially” a $\beta$-Hermite ensemble when $p$ is large enough relative to $n$.

The proof of Theorem 3.4 is different from the standard method of proving large deviations for the empirical distributions of eigenvalues (see, e.g., Ben Arous and Guionnet (1997), Hiai and Petz (1998) and Guionnet (2004)). In fact, reviewing (3.6) and (3.7), we estimate $P(\mu_n \in A)$ for a set $A$ by making a measure transformation such that the underlying $\beta$-Laguerre distribution is changed to a $\beta$-Hermite distribution. After an approximation step similar to that in Theorem 3.1, we use the known result on LDP for $\beta$-Hermite ensembles to complete the proof.

Finally, the order $p \gg n$ in Theorems 3.2 and 3.3 is the best order to make the theorems hold when one considers the case $p$ being much larger than $n$. From the proof of Theorem 3.4, we see the order $p \gg n^2$ is “almost necessary.” Even so, the large deviation principle may still hold with a different rate function and/or a different speed as $p \gg n$ but the condition that $p \gg n^2$ does not hold, we leave it as a future work.
3.4 Proofs of Results in Section 3.2

We start with a concentration inequality on the \( \beta \)-Hermite ensembles.

**Lemma 3.1** Suppose \( \lambda = (\lambda_1, \cdots, \lambda_n) \) has the joint density function \( f_\beta(\lambda) \) as in (3.4). Then, there exists a constant \( C > 0 \) depending on \( \beta \) only such that

\[
P\left( \max_{1 \leq i \leq n} |\lambda_i| \geq \sqrt{n} t \right) \leq C \cdot e^{-\frac{1}{2}nt^2 + Cnt}
\]

for all \( t > 0 \), \( n \geq 2 \) and \( \beta > 0 \).

Ben Arous et al (2001) study the above probability for the case \( \beta = 1 \) in their Lemma 6.3. Our Lemma 3.1 is stronger than theirs when \( t \) is large. In fact, their bound of the above probability is \( e^{-\delta nt^2} \) with some \( \delta \in (0, 1/2) \).

**Proof.** It is easy to see that the order statistic \( \lambda_{(1)} > \cdots > \lambda_{(n)} \) has density function \( h_\beta(\lambda_1, \cdots, \lambda_n) := n! f_\beta(\lambda_1, \cdots, \lambda_n) \) for \( \lambda_1 > \cdots > \lambda_n \). Further, \( \max_{1 \leq i \leq n} |\lambda_i| = |\lambda_{(1)}| \lor |\lambda_{(n)}| \). It follows that

\[
P\left( \max_{1 \leq i \leq n} |\lambda_i| \geq \sqrt{n} t \right) = n! K_n^\beta \cdot \int_{\lambda_1 > \cdots > \lambda_n; |\lambda_1| \lor |\lambda_{n}| \geq \sqrt{n} t} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^{\beta} \cdot e^{-\frac{1}{2} \sum_{i=1}^n \lambda_i^2} d\lambda_1 \cdots d\lambda_n
\]

\[
= \frac{n! K_n^\beta}{(n - 2)! K_{n-2}^\beta} \cdot \int_{\lambda_1 > \cdots > \lambda_{n-1}; |\lambda_1| \lor |\lambda_{n-1}| \geq \sqrt{n} t} \prod_{i=1}^{n-1} |\lambda_i - \lambda_{n-1}|^{\beta} \prod_{j=2}^{n-1} |\lambda_j - \lambda_1|^{\beta} \cdot e^{-\frac{1}{2}(\lambda_1^2 + \lambda_{n-1}^2)} d\lambda_1 d\lambda_{n-1} \cdot g(\lambda_2, \cdots, \lambda_{n-1}) d\lambda_2 \cdots d\lambda_{n-1}
\]

where \( K_n^\beta \) is as in (3.5) and

\[
g(\lambda_2, \cdots, \lambda_{n-1}) = (n - 2)! K_{n-2}^\beta \prod_{2 \leq i < j \leq n-1} |\lambda_i - \lambda_j|^{\beta} \cdot e^{-\frac{1}{2} \sum_{i=2}^{n-1} \lambda_i^2} \tag{3.20}
\]

for \( \lambda_2 > \cdots > \lambda_{n-1} \). Notice

\[
\prod_{i=1}^{n-1} |\lambda_i - \lambda_{n-1}|^{\beta} \cdot \prod_{j=2}^{n-1} |\lambda_j - \lambda_1|^{\beta} \leq (|\lambda_1| + |\lambda_{n-1}|)^{(2n-3)\beta} \leq (2(\lambda_1^2 + \lambda_{n-1}^2))^{(2n-3)\beta/2} \leq 2^{\beta n} (\lambda_1^2 + \lambda_{n-1}^2)^{(2n-3)\beta/2}
\]
for $\lambda_1 > \cdots > \lambda_n$. Further, if $|\lambda_1| \lor |\lambda_n| \geq \sqrt{n}t$, then $\lambda_1^2 + \lambda_n^2 \geq nt^2$. Therefore,

$$P\left( \max_{1 \leq i \leq n} |\lambda_i| \geq \sqrt{n}t \right) \leq n^2 2^{2\beta_n} \frac{K_{n-2}^\beta}{K_{n-2}^\beta} \cdot \int_{x^2+y^2 \geq nt^2} \frac{(x^2 + y^2)^{\frac{1}{2}(2n-3)\beta}}{2} e^{-\frac{1}{2}(x^2+y^2)} \, dx \, dy$$

$$= n^2 2^{2\beta_n} \frac{K_{n-2}^\beta}{K_{n-2}^\beta} \cdot \int_{x^2+y^2 \geq nt^2} (x^2 + y^2)^{\frac{1}{2}(2n-3)\beta} e^{-\frac{1}{2}(x^2+y^2)} \, dx \, dy \quad (3.22)$$

since $g(\lambda_2, \cdots, \lambda_{n-1})$ is a probability density function. By making transform $x = r \cos \theta$ and $y = r \sin \theta$ with $r \geq \sqrt{nt}$ and $\theta \in [0, 2\pi]$, the last integral is equal to

$$2\pi \int_{\sqrt{nt}}^\infty r^{(2n-3)\beta+1} e^{-r^2/2} \, dr = \pi \int_{nt^2}^{\infty} y^{\frac{3}{2}(2n-3)\beta} e^{-y/2} \, dy \quad (3.23)$$

by making another transform $y = r^2$. To compute the last integral, let’s consider $I = \int_b^\infty y^\alpha e^{-y/2} \, dy$ for $b > 0$ and $\alpha > 0$. Use $e^{-y/2} = -2(e^{-y/2})'$ and the integration by parts to have

$$bI \leq \int_b^\infty y^{\alpha+1} e^{-y/2} \, dy = 2b^{\alpha+1} e^{-b/2} + 2(\alpha + 1)I,$$

which implies that

$$I = \int_b^\infty y^\alpha e^{-y/2} \, dy \leq \frac{2}{b - 2\alpha - 2} b^{\alpha+1} e^{-b/2} \quad (3.24)$$

if $\alpha > 0$ and $b > 2\alpha + 2$. Now, suppose $t > \sqrt{4\beta + 4}$, then $nt^2 - (\beta(2n-3) + 2) > \frac{1}{2} nt^2$. By (3.24), the right hand side of (3.23) is bounded by

$$\frac{2\pi}{nt^2 - \beta(2n-3) - 2} (nt^2)^{\frac{3}{2}(2n-3)+1} e^{-nt^2/2} \leq (4\pi)(nt^2)^{\frac{3}{2}(2n-3)} e^{-nt^2/2} \quad (3.25)$$

Now we estimate the term $K_{n-2}^\beta / K_{n-2}^\beta$ appeared in (3.22). By the Stirling formula (see, e.g., p.204 from Ahlfors (1979) or p.368 from Gamelin (2001)),

$$\log \Gamma(z) = z \log z - z - \frac{1}{2} \log 2\pi + \frac{1}{12z} + O\left( \frac{1}{z^3} \right) \quad (3.26)$$
as $x = \text{Re}(z) \to +\infty$. It is easy to check that there exists an absolute constant $C > 0$ such that $\Gamma(1 + x) \geq x^e e^{-Cx}$ for all $x > 0$. Now, recalling $K_n^\beta$ as in (3.5), we have

$$\frac{K_n^\beta}{K_{n-2}^\beta} = \frac{1}{2\pi} \cdot \frac{\Gamma\left(1 + \frac{\beta}{2}(n-1)\right)}{(1 + \frac{\beta}{2}(n-1)) \cdot \Gamma\left(1 + \frac{\beta}{2}n\right)} \leq C\beta^\beta \left(\frac{n}{2}\right)^{\beta(n-1)} e^{\beta n C}.$$ 

Use $n - 1 \geq n/2$ to have

$$\left(\frac{n}{2}\right)^{\beta(n-1)} \leq \left[\left(\frac{2}{\beta}\right)^\beta + 1\right]^n \left(\frac{n}{2}\right)^{\beta(n-1)} \leq e^{C\beta n} \cdot 4^{\beta n - \beta n}$$

since $(n/2)^{-\beta(n-1)} = (2^{n-1}n)^{-\beta n} \leq 4^{\beta n - \beta n}$. Combining the last two inequalities we see that $K_n^\beta/K_{n-2}^\beta \leq C\beta^\beta (\beta^\beta (n-1) - \beta (n-1)) e^{-\beta n C}$.

Combining the last two inequalities we see that

$$\frac{K_n^\beta}{K_{n-2}^\beta} \leq C\beta^\beta \left(\frac{n}{2}\right)^{\beta(n-1)} e^{\beta n C}.$$ 

This together with (3.22) and (3.25) concludes that, for some constant $\gamma$ depending on $\beta$ only,

$$P\left(\max_{1 \leq i \leq n} |\lambda_i| \geq \sqrt{n} t\right) \leq \gamma n^{\gamma} t^{n\gamma} \cdot \exp\left(-\frac{1}{2} nt^2 + \gamma n\right) \leq \gamma \cdot \exp\left(-\frac{1}{2} nt^2 + \gamma n(t + 2)\right) \leq (2\gamma) \cdot \exp\left(-\frac{1}{2} nt^2 + (2\gamma)nt\right)$$

for all $t > \sqrt{4\beta + 4}$, where the inequality $n^\gamma \cdot t^{n\gamma} \leq e^{\gamma nt}$ is used in the second inequality. Note that the last term in (3.27) is increasing in $\gamma > 0$. Set $\gamma' = \gamma + \sqrt{\beta + 1}$, then $2\gamma' nt \geq nt^2$ for all $0 \leq t \leq \sqrt{4\beta + 4}$. It follows that

$$\inf_{0 \leq t \leq \sqrt{4\beta + 4}} \left\{(2\gamma') \cdot \exp\left(-\frac{1}{2} nt^2 + (2\gamma')nt\right)\right\} \geq \inf_{0 \leq t \leq \sqrt{4\beta + 4}} \left\{2\gamma' \cdot e^{nt^2/2}\right\} \geq 1.$$ 

Therefore, (3.27) holds with $C = 2\gamma'$ for all $t \geq 0$, $\beta > 0$ and $n \geq 2$. ■

**Proof of Theorem 3.1.** Let $f_{n,\beta}(\lambda) = f_{n,\beta}(\lambda_1, \cdots, \lambda_n)$ be as in (3.2). Recall $x_i = \sqrt{\frac{p}{2\beta}} \left(\frac{\lambda_i}{p} - \beta\right)$ for $1 \leq i \leq n$ and $\lambda_i = p\beta + \sqrt{2\beta p} x_i$. By (3.2), $x = (x_1, \cdots, x_n)$ has density function

$$\tilde{f}_{n,\beta}(x) = c_n^{\beta p} \prod_{1 \leq i < j \leq n} |x_i - x_j|^\beta \cdot \prod_{i=1}^n \left(1 + \sqrt{\frac{2}{\beta p}} x_i\right)^{\beta(p-1)n} \cdot e^{-\sqrt{\frac{2}{\beta p}} \sum_{i=1}^n x_i}$$
for all $x_i \geq -\sqrt{\frac{2p}{2}}$ with $i = 1, \ldots, n$; $\tilde{f}_{n,\beta}(x) = 0$ otherwise, where
\[
\tilde{c}_n^{\beta,p} = c_n^{\beta,p} \cdot (2\beta p)^{\frac{n(n-1)}{2} + \frac{1}{2}} \cdot e^{-\frac{1}{2}np\beta} \cdot (p\beta)^{\frac{1}{2}n(p-n+1)\beta - n}.
\] (3.28)

Let $\mu_1, \ldots, \mu_n$ have density function $f_\beta(\mu)$ as in (3.4). Then, by (3.9),
\[
\|\mathcal{L}(x_1, \ldots, x_n) - \mathcal{L}(\mu_1, \ldots, \mu_n)\| = \int_{\mathbb{R}^n} |\tilde{f}_{n,\beta}(x) - f_\beta(x)| \, dx_1 \cdots dx_n
\]
\[
= E \left| \frac{\tilde{f}_{n,\beta}(X)}{f_\beta(X)} - 1 \right|
\] (3.29)

where the random vector $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ has density function $f_\beta(x)$ (replacing $\lambda$ and $\lambda_i$ in (3.4) by $x$ and $x_i$ accordingly). Now,
\[
\frac{\tilde{f}_{n,\beta}(X)}{f_\beta(X)} = \frac{\tilde{c}_n^{\beta,p}}{K_n} \prod_{i=1}^n \left( 1 + \sqrt{\frac{2}{\beta p}x_i} \right)^{\frac{2(p-n+1)-1}{2}} \cdot e^{-\sqrt{\frac{2p}{\beta p} \sum_{i=1}^n x_i + \frac{1}{2} \sum_{i=1}^n x_i^2} - \beta x_i} \] (3.30)

for all $x_i \geq -\sqrt{\frac{2p}{2}}$ with $i = 1, \ldots, n$, and it is equal to 0, otherwise.

Recall the two conditions: (i) $n \to \infty$ and $p = p_n \gg n^3$ and (ii) $n$ is fixed and $p \to \infty$. In case (i) we choose constant $t_n > 0$ for all $n \geq 1$ such that
\[
t_n \to \infty \quad \text{and} \quad t_n^4 \cdot \frac{n^3}{p} \to 0
\] (3.31)
as $n \to \infty$. In case (ii) we choose $t_p$ for all $p \geq 1$ satisfying
\[
t_p \to \infty \quad \text{and} \quad t_p^4 \cdot \frac{n^3}{p} \to 0
\] (3.32)
as $p \to \infty$. From now on we will only prove the theorem for case (i). The proof for case (ii) will be carried through by replacing “$t_n$” in (3.31) with “$t_p$” in (3.32) and “$n \to \infty$” with “$p \to \infty$” in the context.

By Lemma 3.1, $P\left( \max_{1 \leq i \leq n} |x_i| \geq \sqrt{n}t_n \right) \leq C \cdot e^{-\frac{1}{2}nt_n^2} \to 0$ as $n \to \infty$. Set
\[
\Omega_n = \left\{ \max_{1 \leq i \leq n} \frac{|x_i|}{\sqrt{n}} \leq t_n \right\}, \quad n \geq 1.
\] (3.33)

Then $P(\Omega_n) \to 1$ as $n \to \infty$. By the Taylor expansion, there exists $\epsilon_0 \in (0, 1)$ such that
\[
\log(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 + u(x) \quad \text{with} \quad |u(x)| \leq |x|^4
\] (3.34)
for all $|x| < \epsilon_0$. Notice $\sqrt{\frac{2}{\beta p}} |x| < \epsilon_0$ on $\Omega_n$ as $n$ is large enough by (3.31). It follows that

$$
\log \prod_{i=1}^{n} \left(1 + \sqrt{\frac{2}{\beta p}} x_i\right)^{\frac{\beta}{2}} (p-n+1) - 1
= \left(\frac{\beta}{2} (p-n+1) - 1\right) \sum_{i=1}^{n} \log \left(1 + \sqrt{\frac{2}{\beta p}} x_i\right)
= \left(\frac{\beta}{2} (p-n+1) - 1\right) \left(\sqrt{\frac{2}{\beta p}} \sum_{i=1}^{n} x_i - \frac{1}{\beta p} \sum_{i=1}^{n} x_i^2\right)
+ \frac{1}{3} \sqrt{\frac{8}{\beta^3 p^3}} \sum_{i=1}^{n} x_i^3 + \sum_{i=1}^{n} \left(\frac{2}{\beta p} x_i\right)^3
$$

(3.35)

on $\Omega_n$ as $n$ is sufficiently large. By writing $\frac{\beta}{2} (p-n+1) - 1 = \frac{\beta}{2} p - \frac{(\beta(n-1)}{2} + 1$, we have

$$
U_{n,1} = \left(\frac{\beta}{2} (p-n+1) - 1\right) \sqrt{\frac{2}{\beta p}} \sum_{i=1}^{n} x_i
= \sqrt{\frac{\beta p}{2}} \sum_{i=1}^{n} x_i + \delta_{n,1} \cdot \left(\frac{n^3}{p}\right)^{1/2} \left(\sum_{i=1}^{n} \frac{x_i}{\sqrt{n}}\right)
$$

(3.36)

with $|\delta_{n,1}| \leq C_{\beta,1}$, where $C_{\beta,1} > 0$ is a constant depending on $\beta$ only. Second,

$$
U_{n,2} := -\left(\frac{\beta}{2} (p-n+1) - 1\right) \cdot \frac{1}{\beta p} \sum_{i=1}^{n} x_i^2
= -\frac{1}{2} \sum_{i=1}^{n} x_i^2 + \delta_{n,2} \cdot \left(\frac{n}{p} \sum_{i=1}^{n} x_i^2\right)
= -\frac{1}{2} \sum_{i=1}^{n} x_i^2 + \delta_{n,2} \cdot \frac{n^3 t_2^2}{p}
$$

(3.37)

on $\Omega_n$ with $|\delta_{n,2}| \leq C_{\beta,2}$, where $C_{\beta,2} > 0$ is a constant depending on $\beta$ only. Now,

$$
U_{n,3} = \left(\frac{\beta}{2} (p-n+1) - 1\right) \cdot \frac{1}{3} \sqrt{\frac{8}{\beta^3 p^3}} \sum_{i=1}^{n} x_i^3
= \delta_{n,3} \cdot \left(\frac{n^3}{p}\right)^{1/2} \sum_{i=1}^{n} \left(\frac{x_i}{\sqrt{n}}\right)^3
$$

(3.38)
with $|\delta_{n,3}| \leq C_{\beta,3}$, where $C_{\beta,3} > 0$ is a constant depending on $\beta$ only. On $\Omega_n$, by (3.34)

$$U_{n,4} := \left(\frac{\beta}{2}(p-n+1) - 1\right) \cdot \left| \sum_{i=1}^{n} u\left(\sqrt{\frac{\beta}{\beta p}} x_i\right) \right| \leq \delta_{n,4} \cdot \left(\frac{1}{p} \sum_{i=1}^{n} |x_i|^4\right) \leq \delta_{n,4} \cdot \frac{n^3 t_n^4}{p} \quad (3.39)$$

with $|\delta_{n,4}| \leq C_{\beta,4}$, where $C_{\beta,4} > 0$ is a constant depending on $\beta$ only. We claim that

$$U_{n,1} = \sqrt{\frac{\beta p}{2}} \sum_{i=1}^{n} x_i + o_P(1), \quad U_{n,2} = -\frac{1}{2} \sum_{i=1}^{n} x_i^2 + o_P(1),$$

$$U_{n,3} = o_P(1), \quad U_{n,4} = o_P(1) \quad (3.40)$$
as $n \to \infty$, where by $Z_n = o_P(1)$ we mean $Z_n \to 0$ in probability as $n \to \infty$.

Looking at (3.37) and (3.39) together with (3.31) and the fact $P(\Omega_n) \to 1$, the claims for $U_{n,2}$ and $U_{n,4}$ in (3.40) are obviously true. By Theorem 1.2 from Dumitriu and Edelman (2006), for each integer $k \geq 1$, $\sum_{i=1}^{n} (x_i/\sqrt{n})^{2k-1}$ converges in distribution to $N(0, \sigma_k^2)$ with $\sigma_k^2 < \infty$ as $n \to \infty$. Reviewing (3.36) and (3.38), from the condition $p \gg n^3$, (3.31) and the fact $P(\Omega_n) \to 1$, we see that the claims for $U_{n,2}$ and $U_{n,4}$ in (3.40) hold true. (If $p_n \equiv p \geq 2$ for all $n \geq 1$ then the claims for $U_{n,1}$ and $U_{n,3}$ are evidently true by (3.36) and (3.38) together with the fact $P(\Omega_n) \to 1$).

Now, from (3.35)-(3.39), we have

$$\log \prod_{i=1}^{n} \left(1 + \sqrt{\frac{2}{\beta p}} x_i\right)^{\frac{\beta}{2}(p-n+1) - 1} = U_{n,1} + U_{n,2} + U_{n,3} + U_{n,4}$$
on $\Omega_n$. Consequently, from (3.30), (3.40) and the fact $P(\Omega_n) \to 1$, we conclude that

$$\frac{\tilde{f}_{n,\beta}(X)}{f_\beta(X)} \cdot \left(\frac{\tilde{c}_{n,\beta}^{\beta p}}{K_n^{\beta}}\right)^{-1} \to 1 \quad (3.41)$$
in probability as $n \to \infty$. Next we show

$$\lim_{n \to \infty} \frac{\tilde{c}_{n,\beta}^{\beta p}}{K_n^{\beta}} = 1. \quad (3.42)$$
Recall (3.3) and (3.5), we know
\[
\frac{c_{n}^{{\beta ,p}}}{K_{n}^{{\beta }}} = (2\pi )^{n/2} \cdot 2^{-\beta np/2} \prod _{j=1}^{n} \frac{1}{\Gamma (\frac{\beta }{2}(p - n + j))} = (2\pi )^{n/2} \cdot 2^{-\beta np/2} \prod _{i=0}^{n-1} \frac{1}{\Gamma (\frac{\beta }{2}(p - i))} \quad (3.43)
\]

From (3.26) we have
\[
\log \Gamma \left( \frac{\beta }{2}(p - i) \right) = \frac{\beta }{2}(p - i) \log \left( \frac{\beta }{2}(p - i) \right) - \frac{1}{2} \log \left( \frac{\beta }{2}(p - i) \right) + \log \sqrt{2\pi} + O \left( \frac{1}{p - i} \right)
\]
\[
= \frac{\beta }{2}(p - i) \log(p - i) + \left( \frac{\beta }{2} \log \frac{\beta }{2} - \frac{\beta }{2} \right)(p - i) - \frac{1}{2} \log(p - i) - \frac{1}{2} \log \frac{\beta }{2} + \log \sqrt{2\pi} + O \left( \frac{1}{p} \right) \quad (3.44)
\]
as \(n \to \infty\) uniformly for all \(i = 0, 1, \cdots, n - 1\). Use the condition \(p \gg n\) and the fact
\[
\log(p - i) = \log p - \frac{i}{p} + O \left( \frac{n^2}{p^3} \right) \quad (3.45)
\]
uniformly for all \(0 \leq i \leq n - 1\) as \(n \to \infty\) to have
\[
- \frac{1}{2} \sum_{i=0}^{n-1} \log(p - i) = - \frac{1}{2} n \log p + O \left( \frac{n^2}{p} \right) \quad (3.46)
\]
as \(n \to \infty\). Moreover,
\[
\frac{(\beta }{2} \log \frac{\beta }{2} - \frac{\beta }{2}) \sum_{i=0}^{n-1} (p - i) = \frac{1}{2} \left( \frac{\beta }{2} \log \frac{\beta }{2} - \frac{\beta }{2} \right)n(2p - n + 1). \quad (3.47)
\]
By (3.45) again,
\[
\frac{\beta }{2} \sum_{i=0}^{n-1} (p - i) \log(p - i) = \left[ \frac{\beta }{2} \sum_{i=0}^{n-1} (p - i)(\log p - \frac{i}{p}) \right] + O \left( \frac{n^3}{p} \right)
\]
\[
= \frac{\beta }{4} (\log p)n(2p - n + 1) - \left[ \frac{\beta }{2} \sum_{i=0}^{n-1} \left( i - \frac{i^2}{p} \right) \right] + O \left( \frac{n^3}{p} \right)
\]
\[
= \frac{\beta }{4} (\log p)n(2p - n + 1) - \frac{\beta }{4} n(n - 1) + O \left( \frac{n^3}{p} \right) \quad (3.47)
\]
as \( n \to \infty \), where we use the fact \( \sum_{i=0}^{n-1} (i - \frac{\beta}{p}) = \frac{1}{2}n(n-1) + O\left(\frac{n^3}{p}\right) \) in the last step. This joint with (3.44), (3.46) and (3.47) leads to

\[
\log \left( \prod_{j=1}^{n} \Gamma\left(\frac{\beta}{2}(p-n+j)\right) \right)^{-1} = - \sum_{i=0}^{n-1} \log \Gamma\left(\frac{\beta}{2}(p-i)\right) = - \frac{\beta}{4} (\log p)n(2p-n+1) - \frac{\beta}{2} (\log \frac{\beta}{2} - 1)np \\
+ \frac{\beta}{4} (\log \frac{\beta}{2})n^2 + \frac{1}{2}n \log p + O\left(\frac{n^3}{p}\right) 
\]

(3.48)
as \( n \to \infty \) under the restriction that \( p \gg n \) only. Consequently,

\[
\log \frac{c_{n,p}}{K_n^\beta} = \frac{n}{2} \log(2\pi) - \frac{\beta np}{2} \log 2 - \sum_{i=0}^{n-1} \log \Gamma\left(\frac{\beta}{2}(p-i)\right) \\
- \frac{\beta np}{2} \log 2 + \left(\frac{1}{2} \log \frac{\beta}{2}\right)n + \frac{1}{2}n \log p - \frac{1}{2} (\log \frac{\beta}{2} - \log \frac{\beta}{2})n(2p-n+1) \\
- \frac{\beta}{4} (\log p)n(2p-n+1) + \frac{\beta}{4} n(n-1) + O\left(\frac{n^3}{p}\right) 
\]

(3.49)
as \( n \to \infty \). Reviewing (3.28),

\[
\log \frac{\bar{c}_{n,p}}{K_n^\beta} = \log \frac{c_{n,p}}{K_n^\beta} + \left(\frac{1}{4}n(n-1)\beta + \frac{1}{2}n\right) \log(p\beta) - \frac{1}{2}np\beta \\
+ \left(\frac{1}{2}n(p-n+1)\beta - n\right) \log(p\beta). 
\]

(3.50)
Combining this with (3.49), by a routine but tedious calculation (see it in Appendix), we have

\[
\log \frac{\bar{c}_{n,p}}{K_n^\beta} = O\left(\frac{n^3}{p}\right) \to 0 
\]

(3.51)
as \( n \to \infty \), which implies (3.42). Finally, by (3.41) and (3.42),

\[
\frac{\tilde{f}_{n,\beta}(X)}{f_\beta(X)} \to 1 
\]
in probability as \( n \to \infty \). Obviously, \( E \frac{\tilde{f}_{n,\beta}(X)}{f_{\beta}(X)} = 1 \) for all \( n \geq 2 \). By a variant of the Scheffé Lemma (see, e.g., Corollary 4.2.4 from Chow and Teicher (1997)), the two facts imply that \( E | \frac{\tilde{f}_{n,\beta}(X)}{f_{\beta}(X)} - 1 | \to 0 \) as \( n \to \infty \). The desired conclusion then follows from (3.29).

\[ \blacksquare \]

**Proof of Proposition 3.1.** Let \( \xi_1, \ldots, \xi_n \) have density function \( f_2(\xi_1, \ldots, \xi_n) \) as in (3.4) with \( \beta = 2 \). Then \( y_1 := \xi_1/\sqrt{2}, \ldots, y_n := \xi_n/\sqrt{2} \) have density function

\[
f(y_1, \ldots, y_n) = \text{Const} \cdot \prod_{1 \leq i < j \leq n} |y_i - y_j|^2 \cdot e^{-\sum_{i=1}^{n} y_i^2}
\]

for \((y_1, \ldots, y_n) \in \mathbb{R}^n\). It is shown by Bornemann (2010) (see also Bianchi et al (2010)) that the two random variables \( \tilde{y}_{\min} := \sqrt{2}n^{1/6}(\xi_{\min} + \sqrt{2n}) = n^{1/6}(\xi_{\min} + 2\sqrt{n}) \) and \( \tilde{y}_{\max} := \sqrt{2}n^{1/6}(\xi_{\max} - \sqrt{2n}) = n^{1/6}(\xi_{\max} - 2\sqrt{n}) \) are asymptotic independent, that is,

\[
P(\tilde{y}_{\min} \in A, \tilde{y}_{\max} \in B) - P(\tilde{y}_{\min} \in A) \cdot P(\tilde{y}_{\max} \in B) \to 0 \tag{3.52}
\]

as \( n \to \infty \) for any Borel sets \( A \) and \( B \). Further, \( \tilde{y}_{\max} \) goes weakly to \( U \) and \( \tilde{y}_{\min} \) goes weakly to \( -V \), where \( U \) are \( V \) are i.i.d. with the distribution function \( F_2(x) \) as in (3.10) (see also Tracy and Widom (1993, 1994)). By the assumptions, \( \lambda = (\lambda_1, \ldots, \lambda_n) \) has density function \( f_{n,2}(\lambda) \) as in (3.2). In (4.32) replacing \( \xi_{\min} \) in the expression of \( \tilde{y}_{\min} \) by \( \sqrt{\frac{p}{2}}(\lambda_{\min}/p - 2) \) and \( \xi_{\max} \) in the expression of \( \tilde{y}_{\max} \) by \( \sqrt{\frac{p}{2}}(\lambda_{\max}/p - 2) \), respectively, we obtain from Theorem 3.1 that

\[
P\left( \frac{\lambda_{\min} - \mu_{n,1}}{\sigma_n} \in A, \frac{\lambda_{\max} - \mu_{n,2}}{\sigma_n} \in B \right) - \,
P\left( \frac{\lambda_{\min} - \mu_{n,1}}{\sigma_n} \in A \right) \cdot \left( \frac{\lambda_{\max} - \mu_{n,2}}{\sigma_n} \in B \right) \to 0 \tag{3.53}
\]

as \( n \to \infty \), where \( \mu_{n,1} = 2p - 4\sqrt{pm} \), \( \mu_{n,2} = 2p + 4\sqrt{pm} \) and \( \sigma_n = 2\sqrt{pm}^{1/6} \). That is, \( (\lambda_{\min} - \mu_{n,1})/\sigma_n \) and \( (\lambda_{\max} - \mu_{n,2})/\sigma_n \) are asymptotic independent.

Finally, using the same argument as in the above, the weak convergence of \( \tilde{y}_{\max} \) to \( U \) and that of \( \tilde{y}_{\min} \) to \( -V \), we obtain that \( \frac{\lambda_{\max} - \mu_{n,2}}{\sigma_n} \) converges weakly to \( U \).
and \( \frac{\lambda_{\min} - \mu_{n,1}}{\sigma_n} \) converges weakly to \(-V\) as \(n \to \infty\). This together with (3.53) gives the desired conclusion. \(\blacksquare\)

**Proof of Corollary 3.1.** Let \(\lambda_1, \cdots, \lambda_n\) be the eigenvalues of \(XX^*\). As mentioned before (3.2), we know \(\lambda = (\lambda_1, \cdots, \lambda_n)\) has density function \(f_{n,\beta}(\lambda)\) as in (3.2) with \(\beta = 2\). Recall \(\mu_{n,1} = 2p - 4\sqrt{pn}, \mu_{n,2} = 2p + 4\sqrt{pn}\) and \(\sigma_n = 2\sqrt{pn^{-1/6}}\) in Proposition 3.1. Since \(\sigma_n \to \infty\) and \(\mu_{n,1}/(2p) \to 1\), by the Slusky lemma, \(\lambda_{\min}/(2p) \to 1\) in probability as \(n \to \infty\). Set \(\delta_n = 4\sqrt{n/p}\).

Write

\[
\sqrt{pn^{1/6}}\left(\frac{\lambda_{\max}}{\lambda_{\min}} - 1 - 4\sqrt{\frac{n}{p}}\right) = \frac{2p}{\sigma_n} \cdot \frac{\lambda_{\max} - (1 + \delta_n)\lambda_{\min}}{\lambda_{\min}} = \frac{2p}{\lambda_{\min}} \cdot \left(\frac{\lambda_{\max} - \mu_{n,2}}{\sigma_n} - (1 + \delta_n)\frac{\lambda_{\min} - \mu_{n,1}}{\sigma_n} + \frac{\mu_{n,2} - (1 + \delta_n)\mu_{n,1}}{\sigma_n}\right).
\]

It is easy to check the last term in the parenthesis is equal to \(8n^{7/6}p^{-1/2} \to 0\) since \(p \gg n^3\). Also, \(1 + \delta_n \to 1\) and \(\frac{2p}{\lambda_{\min}} \to 1\) in probability as \(n \to \infty\). These and Proposition 3.1 conclude that \(\sqrt{pn^{1/6}}\left(\frac{\lambda_{\max}}{\lambda_{\min}} - 1 - 4\sqrt{\frac{n}{p}}\right)\) (3.54)
converges weakly to \(U + V\), where \(U\) and \(V\) are i.i.d. with distribution function \(F_2(x)\) as in (3.10). Now, let \(\alpha_n = 2\sqrt{pn^{1/6}}, \beta_n = 1 + 2\sqrt{\frac{n}{p}}\). Recall \(\kappa_n\) defined in (3.14). Observe

\[
\alpha_n(\kappa_n - \beta_n) = \alpha_n\left(\frac{\lambda_{\max}}{\lambda_{\min}} - \beta_n^2\right) \cdot \left(\sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} + \beta_n\right)^{-1} = \left\{\sqrt{pn^{1/6}}\left(\frac{\lambda_{\max}}{\lambda_{\min}} - 1 - 4\sqrt{\frac{n}{p}}\right) - 4n^{7/6}/\sqrt{p}\right\} \cdot \left(\sqrt{\frac{\lambda_{\max}}{\lambda_{\min}}} + \beta_n\right)^{-1} \cdot 2.
\]

From (3.54), we see that \(\lambda_{\max}/\lambda_{\min} \to 1\) in probability as \(n \to \infty\). Also, \(\beta_n \to 1\) and \(4n^{7/6}/\sqrt{p} \to 0\) since \(p \gg n^3\). The desired conclusion then follows from the Slusky lemma and (3.54). \(\blacksquare\)
Proof of Proposition 3.2. Let $\xi_1, \cdots, \xi_n$ have density function $f_\beta(\xi_1, \cdots, \xi_n)$ as in (3.4). Then $y_i := \sqrt{2/\beta} \xi_i, 1 \leq i \leq n,$ have density function $$f(y_1, \cdots, y_n) = \text{Const} \cdot \prod_{1 \leq i < j \leq n} |y_i - y_j|^{\beta} \cdot e^{-\frac{2}{\beta} \sum_{i=1}^{n} y_i^2}$$ for $(y_1, \cdots, y_n) \in \mathbb{R}^n$. Since $-\xi_{\min}$ and $\xi_{\max}$ have the same distribution, by Theorem 1.1 from Ramírez et al (2011), $$n^{1/6}(2\sqrt{n} + \sqrt{2/\beta} \xi_{\min}) \overset{d}{=} n^{1/6}(2\sqrt{n} - \sqrt{2/\beta} \xi_{\max})$$ converges weakly to the distribution of $\Lambda_0$, which is defined below (3.16). Let $\lambda = (\lambda_1, \cdots, \lambda_n)$ have density function $f_{n, \beta}(\lambda)$ as in (3.2). By Theorem 3.1, $$P\left( g_n(\sqrt{p/2\beta}(\lambda_{\min}/p - \beta)) \leq x \right) - P\left( g_n(\xi_{\min}) \leq x \right) \to 0$$ as $n \to \infty$ for any $x \in \mathbb{R}$ and any sequence of Borel measurable functions $\{g_n(t); t \in \mathbb{R}, n \geq 2\}$. Taking $g_n(t) = n^{1/6}(2\sqrt{n} + \sqrt{2/\beta} t)$ to get $$P\left( \frac{\lambda_{\min} - \mu_n}{\sigma_n} \leq x \right) - P\left( n^{1/6}(2\sqrt{n} + \sqrt{2/\beta} \xi_{\min}) \leq x \right) \to 0$$ where $\mu_n = \beta(p - 2\sqrt{np})$ and $\sigma_n = \beta\sqrt{np}^{-1/6}$. By the earlier conclusion, the last probability goes to $H(x) = P(\Lambda_0 \leq x)$ for all continuous point $x$ of $H(x)$. This leads to the desired conclusion. $\blacksquare$

3.5 Proofs of Results in Section 3.3

This section is divided into two subsections. In Subsection 3.5.1 we prove Theorems 3.2 and 3.3 for the large deviations for the largest and the smallest eigenvalues of the $\beta$-Laguerre ensembles. Subsection 3.5.2 is devoted to the proof of Theorem 3.4 for the large deviations for the empirical distributions of the eigenvalues from the same ensembles.

3.5.1 Proof of Theorem 3.2

Proof of Theorem 3.2. It is easy to check that $I(x) > 0$ for all $x \neq \beta$, $I(\beta) = 0$, $\{I(x) \leq c\}$ is compact for any $c \geq 0$, and $I(x)$ is strictly increasing on $[\beta, \infty)$. 
Now, to prove the theorem, we need to show the following

\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left( \frac{\lambda_{\max}}{p} \in F \right) \leq - \inf_{x \in F} I(x)
\]

(3.55)

\[
\liminf_{n \to \infty} \frac{1}{p} \log P\left( \frac{\lambda_{\max}}{p} \in G \right) \geq - \inf_{x \in G} I(x)
\]

(3.56)

for any closed set \( F \subset \mathbb{R} \) and open set \( G \subset \mathbb{R} \).

The proof of (3.55). Obviously, the joint density function of the order statistics \( \lambda_{\max} = \lambda(1) > \cdots > \lambda(n) \) is \( g(\lambda_1, \cdots, \lambda_n) = n! f_{n, \beta}(\lambda_1, \cdots, \lambda_n) \) for all \( \lambda_1 > \cdots > \lambda_n \). Write

\[
g(\lambda_1, \cdots, \lambda_n)
= \frac{n! e^{\beta, p}}{(n-1)! e^{\beta, p-1}} \left( \lambda_1^{\frac{\beta(p-n+1)-1}{2}} e^{-\frac{1}{2} \lambda_1} \prod_{i=2}^{n} (\lambda_1 - \lambda_i)^{\beta} \right) L_n(\lambda_2, \cdots, \lambda_n)
\]

(3.57)

where

\[
L_n(\lambda_2, \cdots, \lambda_n) = (n-1)! e^{\beta, p-1} \prod_{2 \leq i<j \leq n} |\lambda_i - \lambda_j|^{\beta} \prod_{i=2}^{n} \lambda_i^{\frac{\beta(p-n+1)-1}{2}} e^{-\frac{1}{2} \sum_{i=2}^{n} \lambda_i}
\]

(3.58)

Notice that \( \prod_{i=2}^{n} |\lambda_1 - \lambda_i|^{\beta} \leq \lambda_1^{\beta(n-1)} \) for all \( \lambda_1 > \cdots > \lambda_n \). This gives

\[
B_n \leq \lambda_1^{\beta(n+p-1)-1} e^{-\lambda_1/2}
\]

Thus, from (3.57) we have

\[
P\left( \frac{\lambda_{\max}}{p} \geq x \right) = \int_{px<\lambda_1, \lambda_1>\cdots>\lambda_n>0} g(\lambda_1, \cdots, \lambda_n) \, d\lambda_1 \cdots d\lambda_n
\]

\[
\leq A_n \cdot \int_{px<\lambda_1>\cdots>\lambda_n>0} \lambda_1^{\frac{\beta(n+p-1)-1}{2}} e^{-\lambda_1/2} \, d\lambda_1
\]

\[
\cdot \int_{\lambda_2>\cdots>\lambda_n>0} L_n(\lambda_2, \cdots, \lambda_n) \, d\lambda_2 \cdots d\lambda_n
\]

\[
= A_n \cdot \int_{px<y<\lambda_1>\cdots>\lambda_n>0} y^{\frac{\beta(n+p-1)-1}{2}} e^{-y/2} \, dy
\]

(3.59)
since \( L_n(\lambda_2, \cdots, \lambda_n) \) is a probability density function. We claim, as \( n \) is sufficiently large, the following hold:

\[
\int_{px}^{\infty} y^{\beta (n+p-1)} e^{-y/2} dy \leq \frac{2(px)^{\beta (n+p-1)}}{p(x - \beta) - \beta n + \beta - 2} e^{-px/2} \text{ if } x > \beta; \tag{3.60}
\]

\[
\int_{0}^{px} y^{\beta (p-n+1)} e^{-y/2} dy \leq \frac{2(px)^{\beta (p-n+1)}}{(\beta - x)p - \beta n} e^{-px/2} \text{ if } 0 < x < \beta. \tag{3.61}
\]

In fact, taking \( \alpha = \frac{\beta}{2} (n + p - 1) - 1 \) and \( b = px \) in (3.24) and using the fact \( x > \beta \), we obtain (3.60). To prove (3.61), set \( J = \int_{0}^{b} y^{\alpha - 1} e^{-y/2} dy \) with \( \alpha = \frac{\beta}{2} (p - n + 1) \), \( b = px \) and \( 0 < x < \beta \). By integration by parts,

\[
\alpha J = \int_{0}^{b} (y^{\alpha})' e^{-y/2} dy = b^{\alpha} e^{-b/2} + \frac{1}{2} \int_{0}^{b} y^{\alpha} e^{-y/2} dy \leq b^{\alpha} e^{-b/2} + \frac{b}{2} J.
\]

Solve the inequality to have

\[
J \leq \frac{2}{2\alpha - b} b^{\alpha} e^{-b/2} \leq \frac{2}{(\beta - x)p - \beta n} (px)^{\beta (p-n+1)} e^{-px/2};
\]

which leads to (3.61).

Now we estimate \( A_n \) in (3.57). In fact, by (3.3),

\[
A_n = n \frac{c_{n}^{\beta, p}}{c_{n-1}^{\beta, p - 1}} = \frac{n 2^{-\beta (n+p-1)/2} \Gamma(1 + \frac{\beta}{2})}{\Gamma(1 + \frac{\beta}{2} n) \Gamma(\frac{\beta}{2} p)}. \tag{3.62}
\]

Use the fact \( \Gamma(x + 1) = x \Gamma(x) \) to have

\[
A_n = \frac{2}{\beta} \cdot \frac{2^{-\beta (n+p-1)/2} \Gamma(1 + \frac{\beta}{2})}{\Gamma(\frac{\beta}{2} n) \Gamma(\frac{\beta}{2} p)}.
\]

By (3.26),

\[
\log A_n = -\frac{\beta}{2} (n + p - 1) \log 2 - \log \Gamma\left(\frac{\beta}{2} n\right) - \log \Gamma\left(\frac{\beta}{2} p\right) + o(p)
\]

\[
= -\frac{\beta}{2} (\log 2) p - \frac{\beta}{2} n \log n - \frac{\beta}{2} p \log \left(\frac{\beta}{2} p\right) + o(p)
\]

\[
= -\frac{\beta}{2} p \left(\log 2 - 1 + \log \frac{\beta}{2}\right) - \frac{\beta}{2} n \log n - \frac{\beta}{2} p \log p + o(p) \tag{3.63}
\]
as \( n \to \infty \). Combining (3.59) and (3.60) we have

\[
\log P\left( \frac{\lambda_{\max}}{p} \geq x \right) \leq \frac{\beta}{2} p (\log 2 - 1 + \log 2) - \frac{\beta}{2} n \log n - \frac{\beta}{2} p \log p \\
- \frac{px}{2} + \frac{\beta}{2} (p + n - 1) \log p + \frac{\beta}{2} (p + n - 1) \log x + o(p) \\
= p\left( \frac{\beta}{2} - \frac{\beta}{2} \log \frac{x}{2} + \frac{\beta}{2} \log x \right) - \frac{\beta}{2} n \log n + \frac{\beta}{2} n \log p + o(p)
\]

as \( n \to \infty \) for all \( x > \beta \). Since \(-\frac{\beta}{2} n \log n + \frac{\beta}{2} n \log p = -\frac{\beta}{2} n \log \frac{n}{p} = o(p) \) as \( n \to \infty \), we arrive at

\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left( \frac{\lambda_{\max}}{p} \geq x \right) \leq -(\frac{x}{2} - \frac{\beta}{2} \log \frac{x}{\beta})
\]

(3.65)

for any \( x > \beta \), and hence the same holds for \( x \geq \beta \) since the right hand side of (3.65) is equal to zero when \( x = \beta \).

Now, if \( \frac{1}{p} \lambda_{\max} \leq x \in (0, \beta) \), by the definition of \( f_{n,\beta}(\lambda) \) in (3.2), we see

\[
f_{n,\beta}(\lambda) \leq c_n^{\beta,p}(px)^{\beta n(n-1)/2} \cdot \prod_{i=1}^{n} \frac{\lambda_i^{\beta(p-n+1)-1}}{\beta p} \cdot e^{-\frac{1}{2} \sum_{i=1}^{n} \lambda_i}
\]

where \( c_n^{\beta,p} \) is as in (3.3). It follows that

\[
P\left( \frac{\lambda_{\max}}{p} \leq x \right) \leq c_n^{\beta,p}(px)^{\beta n(n-1)/2} \cdot \left( \int_0^{px} y^{\frac{\beta(p-n+1)-1}{2}} e^{-y/2} dy \right)^n \\
\leq c_n^{\beta,p}(px)^{\beta n(n-1)/2} \cdot \left( \frac{2}{(\beta - x)p - \beta n} (px)^{\frac{\beta(p-n+1)}{2}} e^{-px/2} \right)^n \\
\leq C n \cdot c_n^{\beta,p} \cdot (px)^{\beta np/2} \cdot p^{-n} \cdot e^{-np\beta x/2}
\]

as \( n \) is sufficiently large, where \( C \) is a constant not depending on \( n \) and the second inequality follows from (3.61). Consequently,

\[
\limsup_{n \to \infty} \frac{1}{np} \log P\left( \frac{\lambda_{\max}}{p} \leq x \right) \leq -\frac{x}{2} + \frac{\beta}{2} \log x + \limsup_{n \to \infty} \frac{1}{np} \log \left( p^{\beta np/2} c_n^{\beta,p} \right)
\]

(3.66)
From (3.3) we have
\[
\frac{1}{np} \log \left( p^{\beta np/2} c_{n}^{\beta p} \right) 
\leq \frac{\beta}{2} \log p - \frac{\beta}{2} \log 2 + \frac{1}{np} \log \left( \prod_{j=1}^{n} \Gamma \left( \frac{\beta}{2} (p - n + j) \right) \right)^{-1} 
+ \frac{1}{np} \log \left( \prod_{j=1}^{n} \Gamma (1 + \frac{\beta}{2} j) \right)^{-1} + o(1)
\]  
(3.67)
as \( n \to \infty \). It is known from the paragraph below (3.25) that \( \Gamma (1+x) \geq x^{x}e^{-Cx} \geq 1 \) for \( x \geq e^{C} \), where \( C \) is an universal constant. Then,
\[
\log \prod_{j=1}^{n} \Gamma (1 + \frac{\beta}{2} j) \geq \sum_{j=1}^{n} \left( \frac{\beta}{2} j \right) \log \left( \frac{\beta}{2} j \right) - C \sum_{j=1}^{n} \frac{\beta}{2} j + O(1) 
= \frac{\beta}{2} \sum_{j=2}^{n} j \log j + O(n^2)
\]
as \( n \to \infty \). Notice \( \sum_{j=2}^{n} j \log j \geq \int_{1}^{n} x \log x \, dx = (\frac{1}{2} x^2 \log x - \frac{x^2}{4})|_{1}^{n} = \frac{1}{2} n^2 \log n + o(n^2) \). Hence,
\[
\log \prod_{j=1}^{n} \Gamma (1 + \frac{\beta}{2} j) \geq \frac{\beta}{4} n^2 \log n + O(n^2)
\]
as \( n \to \infty \). From this and (3.48) we get that the sum of the third and fourth terms in (3.67) is bounded by
\[
\frac{1}{np} \left( - \frac{\beta}{4} (\log p) n (2p - n + 1) - \frac{\beta}{2} (\log \frac{\beta}{2} - 1) n p + \frac{\beta}{4} (\log \frac{\beta}{2} n^2 + \frac{1}{2} n \log p - \frac{\beta}{4} n^2 \log n) \right) + o(1)
\]
\[
= - \frac{\beta}{2} \log p - \left( \frac{\beta}{4} \right) n \log \frac{n}{p} - \frac{\beta}{2} (\log \frac{\beta}{2} - 1) + o(1)
\]
\[
= - \frac{\beta}{2} \log p - \frac{\beta}{2} (\log \frac{\beta}{2} - 1) + o(1)
\]
as \( n \to \infty \). Therefore,
\[
\limsup_{n \to \infty} \frac{1}{np} \log \left( p^{\beta np/2} c_{n}^{\beta p} \right) \leq - \frac{\beta}{2} \log p + \frac{\beta}{2}.
\]  
(3.68)
This joint with (3.66) yields that
\[
\limsup_{n \to \infty} \frac{1}{np} \log P\left( \frac{\lambda_{\max}}{p} \leq x \right) \leq -\left( \frac{x - \beta}{2} - \frac{\beta}{2} \log \frac{x}{\beta} \right) < 0
\]
since \(0 < x < \beta\). Therefore,
\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left( \frac{\lambda_{\max}}{p} \leq x \right) = -\infty \quad (3.69)
\]
for all \(0 < x < \beta\). To prove (3.55), without loss of generality, we assume \(F \subset [0, \infty)\) and \(\beta \notin F\). Since \(F\) is closed, then either \(F \subset [0, a]\), \(F \subset [b, \infty)\) or \(F \subset [0, a] \cup [b, \infty)\) for some constants \(a \in F\) and \(b \in F\) with \(0 < a < \beta < b\). Thus (3.55) follows trivially from (3.65) and (3.69).

The proof of (3.56). To prove (3.56), it is enough to show
\[
\liminf_{n \to \infty} \frac{1}{p} \log P\left( \frac{\lambda_{\max}}{p} \in G \right) \geq -I(x) \quad (3.70)
\]
for all \(x \in G\), where \(G\) is an open subset of \(\mathbb{R}\). If \(x < \beta\), then (3.70) holds automatically since \(I(x) = \infty\). If \(x = \beta\), noticing \(I(x) = 0\) if and only if \(x = \beta\), we then know from (3.55) that \(\lambda_{\max} / p \to \beta\) in probability as \(n \to \infty\), thus \(P(\lambda_{\max} / p \in G) \to 1\), hence (3.70) is also true.

Now assume \(x > \beta\). Since \(G\) is open, choose constants \(r, a, b\) with \(\beta < r < a < x < b\) and \((a, b) \subset G\). Recall (3.57) and (3.58). Under the restriction that \(pa < \lambda_1 < pb\) and \(\lambda_n < \cdots < \lambda_2 < pr\), we know
\[
\prod_{i=2}^{n} |\lambda_1 - \lambda_i|^\beta \geq (p(a - r))^{(n-1)\beta} \quad \text{and}
\]
\[
\int_{pa}^{pb} \lambda_1^{\beta(p-n+1)-1} e^{-\lambda_1/2} d\lambda_1 \geq p(b - a)(pa)^{\beta(p-n+1)-1} e^{-pb/2}.
\]
Then, by the same argument as in (3.59), we have

\[ P\left(\frac{\lambda_{\text{max}}}{p} \in G\right) \geq P\left(a < \frac{\lambda_1}{p} < b, \frac{\lambda_2}{p} < r\right) \]

\[ = A_n \cdot \int_{pa}^{pb} d\lambda_1 \cdot \int_{pr>\lambda_2>\cdots>\lambda_n>0} L_n(\lambda_2, \cdots, \lambda_n) d\lambda_2 \cdots d\lambda_n \]

\[ \geq A_n \cdot (p(a-r))^{\beta(n-1)} p(b-a)(pa)^\beta(p-n+1)^{-1} e^{-pb/2} \tilde{P}(\frac{\lambda_2}{p} < r) \]

(3.71)

where \( A_n, B_n \) and \( L_n \) are defined in (3.57) and (3.58), and \( \tilde{P}(\frac{\lambda_2}{p} < r) \) stands for the probability of \( \{\frac{\lambda_2}{p} < r\} \) with the underlying probability distribution having density function \( L_n(\lambda_2, \cdots, \lambda_n) \). Observe from (3.58) that the original beta-Laguerre ensemble with parameter \((n, p_n, \beta)\) becomes \( L_n(\lambda_2, \cdots, \lambda_n) \) with parameter \((n-1, p_n-1, \beta)\). Since \( \lim_{n \to \infty} p_n/n = \infty \), we know that \( p'_n := p_{n+1} - 1 \to \infty \) and \( p'_n/n \to \infty \) as \( n \to \infty \), according to the arguments below (3.70), we have \( \lim_{n \to \infty} \tilde{P}(\frac{\lambda_2}{p} < r) = 1 \) since \( r > \beta \).

Thus,

\[ \frac{1}{p} \log P\left(\frac{\lambda_{\text{max}}}{p} \in G\right) \geq \frac{1}{p} \log A_n + \frac{\beta n \log p}{2} + \frac{\beta}{2} \log p + \frac{\beta}{2} \log a - \frac{b}{2} + o(1) \]

(3.72)

as \( n \to \infty \). By (3.63), the right hand side of the above is equal to

\[ -\frac{\beta}{2} (\log 2 - 1 + \log \frac{\beta}{2}) + \frac{\beta}{2} \log a - \frac{1}{2} b + K_n \]

where

\[ K_n = \left( -\frac{\beta n \log p}{2} - \frac{\beta}{2} \log p \right) + \left( \frac{\beta n \log p}{2} + \frac{\beta}{2} \log p \right) + o(1) \]

\[ = -\frac{\beta n}{2 p} \log \frac{n}{p} + o(1) \to 0 \]

as \( n \to \infty \). Now taking \( \liminf_{n \to \infty} \) for the both sides of the inequality in (3.72), and then letting \( a \uparrow x \) and \( b \downarrow x \), we arrive at

\[ \liminf_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\text{max}}}{p} \in G\right) \geq -\frac{\beta}{2} (\log 2 - 1 + \log \frac{\beta}{2}) + \frac{\beta}{2} \log x - \frac{1}{2} x = -I(x) \]
which gives (3.70) for \( x > \beta \).

**Proof of Theorem 3.3.** It is easy to check that \( I(x) > 0 \) for all \( x \neq \beta \), \( I(\beta) = 0 \), \( \{I(x) \leq c\} \) is compact for any \( c \geq 0 \), and \( I(x) \) is strictly decreasing on \((0, \beta)\). Now, to prove the theorem, we need to show that

\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \in F\right) \leq -\inf_{x \in F} I(x) \quad \text{and} \quad (3.73)
\]

\[
\liminf_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \in G\right) \geq -\inf_{x \in G} I(x) \quad (3.74)
\]

for any closed set \( F \subset \mathbb{R} \) and open set \( G \subset \mathbb{R} \).

The proof of (3.73). Obviously, the joint density function of the order statistics \( \lambda_{\min} = \lambda_{(n)} < \cdots < \lambda_{(1)} \) is \( g(\lambda_1, \cdots, \lambda_n) = n! f_{n, \beta}(\lambda_1, \cdots, \lambda_n) \) for all \( \lambda_1 > \cdots > \lambda_n \). Write

\[
g(\lambda_1, \cdots, \lambda_n) = \frac{n! c_n^{p}}{(n-1)! c_{n-1}^{p}} \left( \frac{\lambda_1^{\frac{p}{2}(p-n+1)-1}}{e^{-\frac{1}{2} \lambda_n}} \prod_{i=1}^{n-1} (\lambda_i - \lambda_n)^\beta \right) L_n(\lambda_1, \cdots, \lambda_{n-1}) \quad (3.75)
\]

where

\[
L_n(\lambda_1, \cdots, \lambda_{n-1}) = (n-1)! c_{n-1}^{p} \prod_{1 \leq i < j \leq n-1} |\lambda_i - \lambda_j|^{\beta} \prod_{i=1}^{n-1} \lambda_i^{\frac{\beta}{2}((p-1)-(n-1)+1)-1} e^{-\frac{1}{2} \sum_{i=1}^{n-1} \lambda_i} \quad (3.76)
\]

Observe that \( \prod_{i=1}^{n-1} |\lambda_n - \lambda_i|^{\beta} \leq \lambda_1^{\beta(n-1)} \) for all \( \lambda_1 > \cdots > \lambda_n > 0 \). This gives

\[
B_n \leq (pM)^{\beta n} \cdot \lambda_1^{\frac{\beta}{2}(p-n+1)-1} e^{-\lambda_n/2}
\]

provided \( \lambda_1 \leq pM \geq 1 \). By Theorem 3.2, for any \( M > \beta \), we know that

\[
P\left(\frac{\lambda_{\max}}{p} \geq M\right) \leq e^{-p\gamma M} \quad (3.77)
\]
as $n$ is sufficiently large, where $\gamma_M = \frac{M - \beta}{4} - \frac{\beta}{4} \log \frac{M}{\beta}$. Thus, for any $0 < x < \beta$, we have

$$P\left(\frac{\lambda_{\min}}{p} \leq x\right) \leq P\left(\frac{\lambda_{\min}}{p} \leq x, \frac{\lambda_{\max}}{p} \leq M\right) + e^{-p\gamma_M}$$

$$\leq 2e^{-p\gamma_M} \lor \int_{p \geq \lambda_1 \lor \cdots \lor \lambda_n \leq p} g(\lambda_1, \cdots, \lambda_n) \, d\lambda_1 \cdots d\lambda_n$$

$$\leq 2e^{-p\gamma_M} \lor (pM)^{\beta n} \cdot A_n \cdot \int_0^{px} \frac{\beta}{n} (p-n+1) - \frac{\beta}{2} \log \frac{2}{\beta} \, e^{-\lambda_n/2} \, d\lambda_n \cdot \int_{\lambda_1 \lor \cdots \lor \lambda_{n-1} > 0} L_n(\lambda_1, \cdots, \lambda_{n-1}) \, d\lambda_1 \cdots d\lambda_{n-1}$$

$$= 2e^{-p\gamma_M} \lor (pM)^{\beta n} \cdot A_n \cdot \int_0^{px} \frac{\beta}{n} (p-n+1) - \frac{\beta}{2} \log \frac{2}{\beta} \, d\lambda_n \quad (3.78)$$

since $L_n(\lambda_1, \cdots, \lambda_{n-1})$ is a probability density function. By (3.61), the last integral is bounded by $\frac{2}{(\beta-x)p-\beta n} (px)^{\beta/(p-n+1)} e^{-px/2}$ as $n$ is sufficiently large. Thus, by (3.63),

$$\log D_n \leq \beta n \log(pM) - \frac{\beta}{2} p \left( \log 2 - 1 + \log \frac{\beta}{2} \right) - \frac{\beta}{2} n \log n - \frac{\beta}{2} p \log p$$

$$+ \log \frac{2}{(\beta-x)p-\beta n} + \frac{\beta}{2} (p-n+1)(\log p + \log x) - \frac{px}{2} + o(p)$$

$$= \eta p - \frac{\beta}{2} \left( \frac{n}{p} \log \frac{n}{p} \right) p + o(p)$$

as $n \to \infty$, where $\eta = -\frac{\beta}{2} \left( \log 2 - 1 + \log \frac{\beta}{2} \right) + \frac{\beta}{2} \log x - \frac{x}{2}$. Since $p \gg n$ it follows that

$$\limsup_{n \to \infty} \frac{1}{p} \log D_n \leq -\frac{\beta}{2} \left( \log 2 - 1 + \log \frac{\beta}{2} \right) + \frac{\beta}{2} \log x - \frac{x}{2}$$

$$= -\left( \frac{x-\beta}{2} - \frac{\beta}{2} \log \frac{x}{\beta} \right).$$

Consequently, by the notation $\gamma_M$ and (3.78) we get

$$\limsup_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \leq x\right) \leq -\left[ \left( \frac{x-\beta}{2} - \frac{\beta}{2} \log \frac{x}{\beta} \right) \land \left( \frac{M-\beta}{4} - \frac{\beta}{4} \log \frac{M}{\beta} \right) \right] \quad (3.79)$$
for any $M > \beta$. Send $M \to \infty$ to have
\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \leq x \right) \leq -I(x)
\tag{3.80}
\]
for any $0 < x < \beta$.

Now, if $\frac{1}{p} \lambda_{\min} \geq x > \beta$ and $\frac{1}{p} \lambda_{\max} \leq M$ for some $M > x$, by the definition of $f_{n, \beta}(\lambda)$ in (3.2), we see
\[
f_{n, \beta}(\lambda) \leq c_{\beta,p}^{\beta}(pM)^{\beta n(n-1)/2} \cdot \prod_{i=1}^{n} \lambda_i^\beta \left(\frac{p}{p-n+1}\right)^{-1} \cdot e^{-\frac{1}{2} \sum_{i=1}^{n} \lambda_i}
\]
where $c_{\beta,p}^{\beta}$ is as in (3.3). It follows that
\[
P\left(\lambda_{\min} \geq x, \lambda_{\max} \leq M \right) \leq c_{\beta,p}^{\beta}(pM)^{\beta n(n-1)/2} \cdot \left(\int_{p-x}^{\infty} y^\beta \left(\frac{p}{p-n+1}\right)^{-1} e^{-y/2} dy \right)^n
\leq c_{\beta,p}^{\beta}(pM)^{\beta n(n-1)/2} \cdot \left((px)^\beta \left(\frac{p}{p-n+1}\right) e^{-px/2}\right)^n
\leq Cn^2 \cdot c_{\beta,p}^{\beta} \cdot (px)^{\beta np/2} \cdot e^{-npx/2}
\]
as $n$ is sufficiently large, where $C = C(x, M) > 0$ is a constant not depending on $n$, and the second inequality follows from (3.24) with $b - 2\alpha - 2 = p(x - \beta) + \beta n - \beta - 2 \to \infty$. This implies
\[
\limsup_{n \to \infty} \frac{1}{np} \log P\left(\frac{\lambda_{\min}}{p} \geq x, \frac{\lambda_{\max}}{p} \leq M \right) \leq -\frac{x}{2} + \beta \log x + \limsup_{n \to \infty} \frac{1}{np} \log \left(p^{\beta np/2} c_{\beta,p}^{\beta}\right).
\tag{3.81}
\]
This joint with (3.68) gives
\[
\limsup_{n \to \infty} \frac{1}{np} \log P\left(\frac{\lambda_{\min}}{p} \geq x, \frac{\lambda_{\max}}{p} \leq M \right) \leq -\left(\frac{x - \beta}{2} - \frac{\beta}{2} \log \frac{x}{\beta}\right) < 0
\]
since $x > \beta$. Therefore, by (3.77),
\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \geq x \right)
\leq \limsup_{n \to \infty} \frac{1}{p} \left[\log P\left(\frac{\lambda_{\min}}{p} \geq x, \frac{\lambda_{\max}}{p} \leq M \right) \vee \log P\left(\frac{\lambda_{\max}}{p} \geq M \right)\right]
= -\infty \vee \left(-\frac{M - \beta}{4} + \frac{\beta}{4} \log \frac{M}{\beta}\right)
\]
for all $M > x > \beta$. By letting $M \to \infty$, we see that
\[
\limsup_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \geq x\right) = -\infty
\] (3.82)
for all $x > \beta$. Since $P(\lambda_{\min} > 0) = 1$, by the same argument as in the paragraph below (3.69), we get (3.73) from (3.80) and (3.82).

The proof of (3.74). In order to prove (3.74), we only need to show
\[
\liminf_{n \to \infty} \frac{1}{p} \log P\left(\frac{\lambda_{\min}}{p} \in G\right) \geq -I(x)
\] (3.83)
for all $x \in G \cap (0, \beta]$, where $G$ is an open subset of $\mathbb{R}$. If $\beta \in G$, since $I(x) = 0$ if and only if $x = \beta$, we then know from (3.73) that $\frac{\lambda_{\min}}{p} \to \beta$ in probability as $n \to \infty$, thus $P(\frac{\lambda_{\min}}{p} \in G) \to 1$, hence (3.83) is true for $x = \beta$.

Now it is enough to prove (3.83) for all $x \in G \cap (0, \beta) \neq \emptyset$. For such $x$, since $G$ is open, choose constants $a, b, r$ with $0 < a < x < b < r < \beta$ and $(a, b) \subset G$. Review (3.75) and (3.76). Under the restriction that $pa < \lambda_n < pb$ and $pr < \lambda_{n-1} < \cdots < \lambda_1$, we have
\[
\prod_{i=1}^{n-1} |\lambda_i - \lambda_n|^\beta \geq (p(r-b))^{\beta(n-1)} \quad \text{and}
\]
\[
\int_{pa}^{pb} \lambda_n^\beta (p-n+1)^{-1} e^{-\lambda_n/2} d\lambda_n \geq p(b-a)(pa)^\beta (p-n+1)^{-1} e^{-pb/2}.
\]
Then, by the same argument as in (3.71), we have
\[
P\left(\frac{\lambda_{\min}}{p} \in G\right)
\]
\[
\geq P\left(a < \frac{\lambda_n}{p} < b, \frac{\lambda_{n-1}}{p} > r\right)
\]
\[
= A_n \cdot \int_{\lambda_1 > \cdots > \lambda_{n-1} > pr} \left( \int_{pa}^{pb} B_n d\lambda_n \right) L_n(\lambda_1, \cdots, \lambda_{n-1}) d\lambda_1 \cdots d\lambda_{n-1}
\]
\[
\geq A_n \cdot (p(r-b))^{\beta(n-1)} \cdot p(b-a) \cdot (pa)^\beta (p-n+1)^{-1} e^{-pb/2} \cdot \tilde{P}\left(\frac{\lambda_{n-1}}{p} > r\right)
\]
where $A_n$, $B_n$ and $L_n$ are defined in (3.75) and (3.76), and $\tilde{P}\left(\frac{\lambda_{n-1}}{p} > r\right)$ stands for the probability of $\left\{\frac{\lambda_{n-1}}{p} > r\right\}$ with the underlying probability distribution having
density function $L_n(\lambda_1, \cdots, \lambda_{n-1})$. Noticing $p'_n := p_{n+1} - 1 \to \infty$ and $p'_n/n \to \infty$ as $n \to \infty$ since $\lim_{n\to\infty} p_n/n = \infty$. Then, by (3.73) and the argument between (3.71) and (3.72), we have $\lim_{n\to\infty} \tilde{P}(\lambda_{n-1}/p > r) = 1$ since $r < \beta$. Thus,

$$\frac{1}{p} \log P \left(\frac{\lambda_{\min}}{p} \in G\right) \geq \frac{1}{p} \log A_n + \frac{\beta n \log p}{2} + \frac{\beta}{2} \log p + \frac{\beta}{2} \log a - \frac{b}{2} + o(1)$$

as $n \to \infty$. By (3.63), the right hand side of the above is equal to

$$-\frac{\beta}{2} (\log 2 - 1 + \log \frac{\beta}{2}) + \frac{\beta}{2} \log a - \frac{1}{2} b + K_n$$

where

$$K_n = \left( -\frac{\beta}{2} n \log n - \frac{\beta}{2} \log p \right) + \left( \frac{\beta}{2} n \log p \right) + o(1)$$

as $n \to \infty$. Now taking $\liminf_{n\to\infty}$ for the both sides of (3.84), and then letting $a \uparrow x$ and $b \downarrow x$, we arrive at

$$\liminf_{n\to\infty} \frac{1}{p} \log P \left(\frac{\lambda_{\min}}{p} \in G\right) \geq -\frac{\beta}{2} (\log 2 - 1 + \log \frac{\beta}{2}) + \frac{\beta}{2} \log x - \frac{1}{2} x = -I(x)$$

for all $x \in G \cap (0, \beta)$, which concludes (3.83). ■

3.5.2 Proof of Theorem 3.4

To prove Theorem 3.4 next, we need to review some terminology. Let $\mathcal{M}(\mathbb{R})$ be the collection of the Borel probability measures defined on $\mathbb{R}$ associated with the standard weak topology, that is, $\mu_n$ converges to $\mu$ weakly as $n \to \infty$ if and only if

$$\lim_{n\to\infty} \int f(x) \mu_n(dx) = \int f(x) \mu(dx)$$

for every bounded and continuous function $f(x)$ defined on $\mathbb{R}$, where $\{\mu, \mu_n; n \geq 1\} \subset \mathcal{M}(\mathbb{R})$. For further reference, see, e.g., chapter 11 from Dudley (2002). When we mention open and closed sets in $\mathcal{M}(\mathbb{R})$ in the following, the corresponding topology is the weak topology.
Proof of Theorem 3.4. By Theorem 1.3 from Ben Arous and Guionnet (1997), \( I_\beta(\nu) \) is a good rate function, that is, \( I_\beta(\nu) \geq 0 \) for all \( \nu \in \mathcal{M}(\mathbb{R}) \) and \( \{ \nu \in \mathcal{M}(\mathbb{R}); I_\beta(\nu) \leq l \} \) is compact under the weak topology for any \( l \geq 0 \). So we only need to show

\[
\limsup_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in F) \leq -\inf_{\nu \in F} I_\beta(\nu) \tag{3.85}
\]

for any closed set \( F \subset \mathcal{M}(\mathbb{R}) \) and

\[
\liminf_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in G) \geq -\inf_{\nu \in G} I_\beta(\nu) \tag{3.86}
\]

for any open set \( G \subset \mathcal{M}(\mathbb{R}) \).

The proof of (3.85). Define

\[
E_n(\epsilon) = \left\{ \max_{1 \leq i \leq n} \left| \frac{\lambda_i}{p} - \beta \right| < \sqrt{2}\beta \epsilon \right\} \tag{3.87}
\]

for \( 0 < \epsilon < \frac{1}{4}(\sqrt{\beta} \wedge 1) \). By Theorems 3.2 and 3.3, there exists a constant \( \delta = \delta(\epsilon) > 0 \) such that

\[
P(E_n(\epsilon)^c) \leq P\left( \frac{\lambda_{\max}}{p} \geq \beta + \sqrt{2}\beta \epsilon \right) + P\left( \frac{\lambda_{\min}}{p} \leq \beta - \sqrt{2}\beta \epsilon \right) \leq 2e^{-p\delta} \tag{3.88}
\]
as \( n \) is sufficiently large. By (3.2),

\[
P(\mu_n \in F)
\leq 2e^{-p\delta} + P(\{ \mu_n \in F \} \cap E_n(\epsilon))
= 2e^{-p\delta} + c_n^{\beta,p} \cdot \int_{\{ \mu_n \in F \} \cap E_n(\epsilon)} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^\beta \cdot \prod_{i=1}^n \lambda_i^{\frac{2}{p}(p-n+1)-1} \cdot e^{-\frac{1}{2} \sum_{i=1}^n \lambda_i} d\lambda_1 \cdots d\lambda_n.
\]

Since \( x_i = \sqrt{\frac{p}{2\beta}}(\frac{\lambda_i}{p} - \beta) \). Then \( \lambda_i = p(\beta + \sqrt{\frac{2\beta}{p}} x_i) \) with \( x_i > -\sqrt{\beta p/2} \) for \( 1 \leq i \leq n \). It follows that

\[
P(\mu_n \in F)
\leq 2e^{-p\delta} + C_n^{\beta,p} \cdot \int_{\{ \mu_n \in F \} \cap E_n(\epsilon)} \prod_{1 \leq i < j \leq n} |x_i - x_j|^\beta \cdot \prod_{i=1}^n \left( 1 + \sqrt{\frac{2}{p\beta}} x_i \right)^{\frac{2}{p}(p-n+1)-1} \cdot e^{-\sqrt{\frac{p\beta}{2}} \sum_{i=1}^n x_i} dx_1 \cdots dx_n \tag{3.89}
\]
where
\[ E_n(\epsilon)' = \left\{ \max_{1 \leq i \leq n} |x_i| < \epsilon \right\} \] and
\[ C_n^{\beta p} = c_n^{\beta p} \cdot (2\beta p)^{(n-1)\beta/4} \cdot (p\beta)^{\frac{2n(p-n+1)-n}{p}} \cdot e^{-np\beta/2} \cdot (2\beta p)^{n/2}. \] (3.91)

By inequality \( \log(1 + x) \leq x - \frac{x^2}{2} + \frac{x^3}{3} \) for all \( x > -1 \), we have
\[
\log \prod_{i=1}^{n} \left( 1 + \sqrt[2]{\frac{2}{p\beta x_i}} \right)^{\frac{\beta}{2}(p-n+1)-1} 
\leq \left( \frac{\beta}{2}(p-n+1) - 1 \right) \left( \sqrt[2]{\frac{2}{p\beta}} \sum_{i=1}^{n} x_i - \frac{1}{\beta p} \sum_{i=1}^{n} x_i^2 + \frac{2\sqrt{2}}{3\beta^{3/2}p^{3/2}} \sum_{i=1}^{n} x_i^3 \right)
\]
for \( x_i > -\sqrt{\beta p/2} \) with \( i = 1, \cdots, n \). Now, on \( E_n(\epsilon)' \) we have \( \sum_{i=1}^{n} |x_i| \leq n\sqrt{p\epsilon} \), it follows that
\[
U_n := \left( \frac{\beta}{2}(p-n+1) - 1 \right) \left( \sqrt[2]{\frac{2}{p\beta}} \sum_{i=1}^{n} x_i + \epsilon \cdot O(n^2) \right) \] (3.92)
as \( n \to \infty \). Similarly, on \( E_n(\epsilon)' \) we have \( \sum_{i=1}^{n} |x_i|^2 \leq np\epsilon^2 \), which leads to
\[
V_n := -\left( \frac{\beta}{2}(p-n+1) - 1 \right) \cdot \frac{1}{\beta p} \sum_{i=1}^{n} x_i^2 = -\frac{1}{2} \sum_{i=1}^{n} x_i^2 + \epsilon \cdot O(n^2) \] (3.93)
as \( n \to \infty \) since \( 0 < \epsilon < 1 \). By the same argument, on \( E_n(\epsilon)' \) we have \( \sum_{i=1}^{n} |x_i|^3 \leq \sqrt{p\epsilon} \sum_{i=1}^{n} |x_i|^2 \), hence,
\[
W_n := \left( \frac{\beta}{2}(p-n+1) - 1 \right) \left( \frac{2\sqrt{2}}{3\beta^{3/2}p^{3/2}} \sum_{i=1}^{n} |x_i| \right) \leq \frac{\epsilon}{\sqrt{\beta}} \sum_{i=1}^{n} x_i^2. \] (3.94)

Combining all the above we get
\[
\log \prod_{i=1}^{n} \left( 1 + \frac{x_i}{\sqrt[2]{\beta p}} \right)^{\frac{\beta}{2}(p-n+1)-1} \leq U_n + V_n + W_n \leq \sqrt[2]{\frac{\beta}{p}} \sum_{i=1}^{n} x_i - \frac{\alpha_\epsilon}{2} \sum_{i=1}^{n} x_i^2 + \epsilon \cdot O(n^2)
\]
on \( E_n(\epsilon)' \) as \( n \to \infty \), where
\[
\alpha_\epsilon := 1 - \frac{2\epsilon}{\sqrt{\beta}} > 0 \] (3.95)
for $0 < \epsilon < \frac{1}{4}(\sqrt{\beta} \wedge 1)$. From (3.89) we see that

$$P(\mu_n \in F)$$

$$\leq 2e^{-p\delta} + C_n^{\beta,p} \cdot \exp\{\epsilon \cdot O(n^2)\} \cdot \int_{\mu_n \in F} \prod_{1 \leq i < j \leq n} |x_i - x_j|^\beta$$

$$\cdot e^{-\frac{\alpha}{\pi} \sum_{i=1}^{n} x_i^2} \, dx_1 \cdots dx_n$$

(3.96)

as $n$ is sufficiently large. Let $\lambda_1, \cdots, \lambda_n$ have the density function $f_\beta(\lambda_1, \cdots, \lambda_n)$ as in (3.4). Set $y_i = \lambda_i / \sqrt{\alpha \epsilon}$ for $i = 1, \cdots, n$. We know that $(y_1, \cdots, y_n)$ has density

$$h_\beta(y_1, \cdots, y_n) := \alpha \epsilon \beta^{-\frac{\beta}{\alpha \epsilon}} \frac{n^{(n-1)+\frac{\beta}{2}} K_n^\beta \prod_{1 \leq i < j \leq n} |y_i - y_j|^\beta \cdot e^{-\frac{\alpha}{\pi} \sum_{i=1}^{n} y_i^2}}$$

(3.97)

for $(y_1, \cdots, y_n) \in \mathbb{R}^n$. By Corollary 5.1 from Ben Arous and Guionnet (1997) (taking $a = \alpha \epsilon$, $V(x) = \alpha \epsilon x^2$ and $f(x) \equiv 0$) that $\nu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{y_i/\sqrt{n}}$ satisfies the LDP with speed $\{n^2; n \geq 1\}$ and rate function $\alpha \epsilon \cdot I_{\beta/\alpha \epsilon}(\nu)$ where

$$I_b(\nu) = \frac{1}{2} \int_{\mathbb{R}^2} g_b(x,y) \nu(dx) \nu(dy) + \frac{b}{4} \log \frac{b}{2} - \frac{3}{8} b$$

(3.98)

for any $b > 0$ and

$$g_b(x,y) = \begin{cases} \frac{1}{2}(x^2 + y^2) - b \log |x - y|, & \text{if } x \neq y; \\ +\infty, & \text{otherwise.} \end{cases}$$

(3.99)

We see from (3.96) that

$$P(\mu_n \in F)$$

$$\leq 2e^{-p\delta} + \frac{C_n^{\beta,p}}{K_n^\beta} \cdot \exp\{\epsilon \cdot O(n^2)\} \cdot \alpha \epsilon^{-\frac{\beta}{\alpha \epsilon}} \prod_{i=1}^{n} x_i^2 \cdot P(\nu_n \in F)$$

(3.100)

as $n \to \infty$. It follows that

$$\limsup_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in F)$$

$$\leq \limsup_{n \to \infty} \frac{1}{n^2} \log \left(\frac{C_n^{\beta,p}}{K_n^\beta} \cdot \inf_{\nu \in F} \left\{ \alpha \epsilon \cdot I_{\beta/\alpha \epsilon}(\nu) \right\} - \frac{\beta}{4} \log \alpha \epsilon + \epsilon O(1)\right)$$

(3.101)

where the condition $p \gg n^2$ is used in the inequality. For the constant and rate function above, we have the following facts.
**Lemma 3.2** If \( p = p_n \gg n^2 \), then \( \log \frac{C_n^{\beta,p}}{K_n^\beta} = O(n) \) as \( n \to \infty \).

**Lemma 3.3** Let \( I_b(\nu) \) be defined as in (3.98). Let \( A \) be a set of Borel probability measures on \( \mathbb{R} \). Set \( J_s(A) = \inf_{\nu \in A} \{ s \cdot I_b/\nu(\nu) \} \) for all \( s > 0 \). If \( \{ s_n > 0; n \geq 1 \} \) is a sequence with \( \lim_{n \to \infty} s_n = s \in (0, \infty) \), then \( \lim_{n \to \infty} J_{s_n}(A) = J_s(A) \).

The proofs of the two lemmas will be given at the end of this section. Let’s continue now. From (3.101) and Lemma 3.2,

\[
\limsup_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in F) \leq - \inf_{\nu \in F} \left\{ \alpha_{\epsilon} \cdot I_\beta/\alpha_{\epsilon}(\nu) \right\} - \frac{\beta}{4} \log \alpha_{\epsilon} + \epsilon \cdot O(1) \tag{3.102}
\]

for any \( \epsilon \in (0, \frac{1}{4}(\sqrt{\beta} \wedge 1)) \). Now passing \( \epsilon \downarrow 0 \), we know from (3.95) that \( \alpha_{\epsilon} \to 1 \), hence it follows from Lemma 3.3 that

\[
\limsup_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in F) \leq - \inf_{\nu \in F} I_\beta(\nu). \tag{3.103}
\]

The proof of (3.86). First, by the Taylor expansion, there exists \( \epsilon_0 \in (0, 1) \) such that

\[
\log(1 + x) \geq x - \frac{1}{2} x^2 - |x|^3 \tag{3.104}
\]

for all \( |x| < \epsilon_0 \). Without loss of generality, assume

\[
0 < \epsilon_0 < \frac{1}{4}(\sqrt{\beta} \wedge 1). \tag{3.105}
\]

Review (3.87) and (3.89) to have

\[
P(\mu_n \in G) \geq P(\{ \mu_n \in G \} \cap E_n(\epsilon)) \tag{3.106}
\]

\[
= C_n^{\beta,p} \cdot \int_{\{ \mu_n \in G \} \cap E_n(\epsilon)} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^{\beta} \cdot \prod_{i=1}^{n} |\lambda_i|^{\frac{2}{2}(p-n+1)-1} \cdot e^{-\frac{1}{2} \sum_{i=1}^{n} \lambda_i} \, d\lambda_1 \cdots d\lambda_n \tag{3.107}
\]

\[
= C_n^{\beta,p} \cdot \int_{\{ \mu_n \in G \} \cap E_n(\epsilon)} \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j|^{\beta} \cdot \prod_{i=1}^{n} \left( 1 + \sqrt{\frac{2}{\beta p} x_i} \right)^{\frac{2}{2}(p-n+1)-1} \cdot e^{-\sqrt{\frac{2}{\beta p} \sum_{i=1}^{n} x_i}} \, dx_1 \cdots dx_n
\]
where $E_n(\epsilon)'$ is defined in (3.90) and $C_n^{\beta,p}$ in (3.91). Now, by the inequality in (3.104), on $E_n(\epsilon)'$, we have

$$\log n \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta \sqrt{p}}\right)^{\frac{\beta}{2} (p-n+1) - 1} \geq \left(\frac{\beta}{2} (p-n+1) - 1\right) \left(\sqrt{\frac{2}{\beta p}} \sum_{i=1}^{n} x_i - \frac{1}{\beta p} \sum_{i=1}^{n} x_i^2 - \frac{2\sqrt{2}}{\beta^3 p^{3/2}} \sum_{i=1}^{n} |x_i|^3\right)$$

$$= U_n + V_n - 3W_n$$

where $U_n, V_n$ and $W_n$ are defined in (3.92), (3.93) and (3.94), respectively. Thus,

$$\log n \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta \sqrt{p}}\right)^{\frac{\beta}{2} (p-n+1) - 1} \geq \sqrt{\frac{\beta p}{2}} \sum_{i=1}^{n} x_i - \frac{\gamma \epsilon}{2} \sum_{i=1}^{n} x_i^2 + \epsilon \cdot O(n^2)$$

on $E_n(\epsilon)'$ as $n \to \infty$, where

$$\gamma \epsilon := 1 + \frac{6\epsilon}{\sqrt{\beta}} > 0$$

(3.108)

for any $0 < \epsilon < \epsilon_0$. This joint with (3.107) yields that

$$P(\mu_n \in G) \geq C_n^{\beta,p} \cdot e^{O(n^2)} \cdot \int_{\{\mu_n \in G\} \cap E_n(\epsilon)'} \prod_{1 \leq i < j \leq n} |x_i - x_j|^{\beta} \cdot e^{-\frac{2\beta}{2} \sum_{i=1}^{n} x_i^2} dx_1 \cdots dx_n$$

$$= C_n^{\beta,p} \cdot \frac{\gamma \epsilon}{K_n^\beta} \cdot e^{O(n^2)} \cdot \gamma \epsilon^{\frac{n(n-1)}{2} - \frac{\beta}{2}} \cdot P_1(\{\mu_n \in G\} \cap E_n(\epsilon)')$$

by the same arguments as those in (3.97) and (3.100), where $P_1$ stands for the probability such that $x = (x_1, \cdots, x_n)$ appearing in the definitions of $\mu_n$ and $E_n(\epsilon)'$ has the probability density function

$$\tilde{h}_\beta(x_1, \cdots, x_n) := \gamma \epsilon \cdot \frac{2^\frac{n(n-1)}{2} + \frac{n}{2}}{K_n^\beta} \prod_{1 \leq i < j \leq n} |x_i - x_j|^{\beta} \cdot e^{-\frac{\gamma \epsilon}{2} \sum_{i=1}^{n} x_i^2}$$

for $(x_1, \cdots, x_n) \in \mathbb{R}^n$. From Lemma 3.2, we get

$$\liminf_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in G) \geq \liminf_{n \to \infty} \frac{1}{n^2} \log P_1(\{\mu_n \in G\} \cap E_n(\epsilon)') + \epsilon \cdot O(1) - \frac{\beta}{4} \log \gamma \epsilon$$

(3.109)
for any $0 < \epsilon < \epsilon_0$. Denote $J_\epsilon(G) := \inf_{\nu \in G} \{ \gamma_\epsilon \cdot I_{\beta/\gamma}(\nu) \}$. We claim that (3.109) implies

$$\liminf_{n \to \infty} \frac{1}{n^2} \log P(\mu_n \in G) \geq -J_\epsilon(G) - \epsilon + \epsilon \cdot O(1) - \frac{\beta}{4} \log \gamma_\epsilon \quad (3.110)$$

for all $0 < \epsilon < \epsilon_0$. In fact, the above is trivially true if $J_\epsilon(G) = \infty$. Assume now $J_\epsilon(G) < \infty$. Then, by the LDP discussion between (3.97) and (3.100), we know $P_1(\mu_n \in G) \geq \exp\{-n^2(J_\epsilon(G) + \epsilon)\}$ as $n$ is sufficiently large. From Lemma 3.1,

$$P_1(E_n(\epsilon)^c) = P_1\left( \max_{1 \leq i \leq n} |x_i| \geq \sqrt{p} \epsilon \right) = P\left( \max_{1 \leq i \leq n} |\lambda_i| \geq \sqrt{n}(\sqrt{p/n} \gamma_\epsilon^{1/2}) \right) \leq e^{-\frac{p \gamma_\epsilon^2}{3}}$$

as $n$ is sufficiently large since $p \gg n$, where $(\lambda_1, \cdots, \lambda_n) := \gamma_\epsilon^{1/2} \cdot (x_1, \cdots, x_n)$ has the joint probability density function $f_\beta(\lambda_1, \cdots, \lambda_n)$ as in (3.4). Hence,

$$P_1(\{ \mu_n \in G \} \cap E_n(\epsilon)^c) \geq P_1(\mu_n \in G) - P_1(E_n(\epsilon)^c) \geq e^{-n^2(J_\epsilon(G) + \epsilon)} - e^{-\frac{p \gamma_\epsilon^2}{3}} = e^{-n^2(J_\epsilon(G) + \epsilon)}(1 + o(1))$$

as $n \to \infty$ by the condition $p \gg n^2$. This and (3.109) lead to (3.110). Finally, letting $\epsilon \downarrow 0$ in (3.110), we have (3.86) from Lemma 3.3. \[\Box\]

**Proof of Lemma 3.2.** Review (3.91). Notice

$$\frac{C_n^\beta p}{K_n^\beta} = \frac{e_n^{\beta p}}{K_n^\beta} \cdot (2\beta p)^{n(p-1)/2} \cdot (p/\beta)^{\gamma_n(p-n+1) - n} \cdot e^{-np/2} \cdot (2\beta p)^{n/2}$$

$$= 2^{-\frac{\beta}{2}np} \cdot (2\pi)^{n/2} \cdot \beta^{\gamma_n(p-n+1) - n} \cdot p^{\beta n(2p-n+1)/4-(n/2)} \cdot e^{-\beta np/2} \cdot (2\beta)^{\gamma_n(p-n+1)/4+(n/2)} \cdot B_n \cdot \left( \prod_{j=1}^{n} \Gamma\left(\frac{\beta}{2}(p-n+j)\right) \right)^{-1}. \quad (3.111)$$
Observe

\[
\log D_n = -\frac{\beta}{2} (\log 2)np + \frac{n}{2} \log(2\pi) + \frac{\beta}{2} (\log \beta)n(p - n + 1) - n \log \beta \\
+ \frac{\beta}{4} n(2p - n + 1) \log p - \frac{n}{2} \log p - \frac{\beta}{2} np + \left( \frac{\beta}{4} n(n - 1) + \frac{n}{2} \right) \log(2\beta)
\]

\[
= \frac{\beta}{4} (\log p)n(2p - n + 1) + \frac{\beta}{2} (\log \beta)n(p - n) - \frac{\beta}{2} (1 + \log 2)np - \frac{n}{2} \log p \\
+ \frac{\beta \log(2\beta)}{4} n^2 + O(n)
\]

\[
= \frac{\beta}{4} (\log p)n(2p - n + 1) - \frac{\beta}{4} (\log \beta)n^2 + \frac{\beta}{2} (\log \beta - 1 - \log 2)np \\
- \frac{n}{2} \log p + O(n)
\]
as \(n \to \infty\). Joint this and (3.48) with (3.111) we conclude

\[
\log \frac{C^\beta_p n}{K_n} = O\left(n + \frac{n^3}{p}\right) = O(n)
\]
as \(n \to \infty\) since \(p \gg n^2\). ☐

**Proof of Lemma 3.3.** Recalling the expression of \(I_b(\nu)\) in (3.98) and (3.99), to prove the lemma, it is enough to show that

\[
H_n := \inf_{\nu \in A} \int_{\mathbb{R}^2} \left(a_n(x^2 + y^2) - \log |x - y|\right) v(dx)v(dy)
\]

\[
\to H := \inf_{\nu \in A} \int_{\mathbb{R}^2} \left(a(x^2 + y^2) - \log |x - y|\right) v(dx)v(dy) \quad (3.112)
\]
as \(n \to \infty\), where \(\{a_n > 0; \ n \geq 1\}\) is a sequence with \(\lim_{n \to \infty} a_n = a \in (0, \infty)\).

We first claim that

\[
t(x^2 + y^2) - \log |x - y| + C \geq \frac{t}{2} (x^2 + y^2) \quad (3.113)
\]
for any \(t > 0, \ x > 0 \) and \(y > 0\), where \(C = C_t = -\frac{1}{2} \cdot \inf_{u > 0} \{tu - \log u\}\) is finite. In fact,

\[
\frac{t}{2} (x^2 + y^2) - \log |x - y| \geq \frac{t}{2} (x - y)^2 - \frac{1}{2} \log(x - y)^2 \geq -C,
\]
which yields (3.113). The inequality in (3.113) implies that
\[
\int_{\mathbb{R}^2} \left( t(x^2 + y^2) - \log |x - y| \right) v(dx)v(dy) \text{ is finite if and only if }
\int_{\mathbb{R}} x^2 v(dx) < \infty \quad \text{and} \quad \int_{\mathbb{R}^2} \log |x - y| v(dx)v(dy) \text{ is finite.} \tag{3.114}
\]

The inequality in (3.113) also says that the sequence \( \{H, H_n; n \geq 1\} \) is bounded below.

Now, if \( H < \infty \), take any \( \nu \in A \) such that \( \int_{\mathbb{R}^2} (a(x^2+y^2) - \log |x-y|) v(dx)v(dy) < \infty \). By (3.114),
\[
H_n \leq a_n \int_{\mathbb{R}^2} (x^2 + y^2)v(dx)v(dy) - \int_{\mathbb{R}^2} \log |x - y| v(dx)v(dy) < \infty
\]
for all \( n \geq 1 \). Passing \( n \to \infty \) and taking the infimum over all \( \nu \in A \) for both sides, we get
\[
\limsup_{n \to \infty} H_n \leq H. \tag{3.115}
\]

Obviously, the above is also true if \( H = \infty \).

On the other hand, if \( H < \infty \), by (3.114), \( H_n < \infty \) for all \( n \geq 1 \). For any \( n \geq 1 \), take \( \nu_n \in A \) with
\[
H_n + \frac{1}{n} \geq \int_{\mathbb{R}^2} (a_n(x^2 + y^2) - \log |x - y|) v_n(dx)v_n(dy) \tag{3.116}
\geq \int_{\mathbb{R}^2} (2\delta(x^2 + y^2) - \log |x - y|) v_n(dx)v_n(dy)
\geq \delta \int_{\mathbb{R}^2} (x^2 + y^2)v_n(dx)v_n(dy) - C
\]
where \( 2\delta := \inf\{a_n; n \geq 1\} \) and the last inequality follows from (3.113) with \( C = C_\delta \). This and (3.115) imply that
\[
M := \sup_{n \geq 1} \int_{\mathbb{R}^2} (x^2 + y^2)v_n(dx)v_n(dy) < \infty.
\]

Therefore, from (3.116),
\[
H_n + \frac{1}{n} \geq \int_{\mathbb{R}^2} (a(x^2 + y^2) - \log |x - y|) v_n(dx)v_n(dy) - M|a_n - a|
\geq H - M|a_n - a|.
\]
Letting $n \to \infty$, we have $\lim \inf_{n \to \infty} H_n \geq H$. Further, (3.114) says that $H = \infty$ if and only if $H_n = \infty$ for all $n \geq 1$. Thus, $\lim \inf_{n \to \infty} H_n \geq H$ also holds if $H = \infty$. These and (3.115) prove (3.112). ■

3.6 Appendix

In this part, we verify the validity of (3.51).

Proof of (3.51). Set

$$A = -\frac{\beta np}{2} \log 2 + \left(\frac{1}{2} \log \frac{\beta}{2}\right)n + \frac{1}{2} n \log p - \frac{1}{2} \left(\frac{\beta}{2} \log \frac{\beta}{2} - \frac{\beta}{2}\right)n(2p - n + 1)$$
$$\quad - \frac{\beta}{4}(\log p)n(2p - n + 1) + \frac{\beta}{4}n(n - 1);$$

$$B = \left(\frac{1}{4}n(n - 1)\beta + \frac{1}{2}n\right) \log(2\beta p) - \frac{1}{2} np\beta + \left(\frac{1}{2}n(p - n + 1)\beta - n\right) \log(p\beta).$$

To show (3.51), by (3.49) and (3.50), we need to check that $A + B = 0$. First, separate the terms with $\log p$ from others in the expression of $B$ and then sort out for $\log \beta$ to have

$$B = \left(\frac{1}{4}n(n - 1)\beta + \frac{1}{2}n\right) \log p + \left(\frac{1}{4}n(n - 1)\beta + \frac{1}{2}n\right) \log(2\beta) +$$
$$\quad \left(\frac{1}{2}n(p - n + 1)\beta - n\right) \log p + \left(\frac{1}{2}n(p - n + 1)\beta - n\right) \log \beta - \frac{1}{2} np\beta$$
$$= \left(\frac{\beta}{2} pn \log p - \left(\frac{1}{4}n(n - 1)\beta + \frac{1}{2}n\right) \log p\right) + B'$$

where

$$B' = \frac{(\log \beta)n}{4}((2p - n + 1)\beta - 2) + \left(\frac{1}{4}n(n - 1)\beta + \frac{1}{2}n\right) \log 2 - \frac{1}{2} np\beta.$$

Comparing the two coefficients of $\log p$ in $A$ and $B$, we find that their sum is identical to 0. Thus, by setting

$$A' = -\frac{\beta np}{2} \log 2 + \left(\frac{1}{2} \log \frac{\beta}{2}\right)n - \frac{1}{2} \left(\frac{\beta}{2} \log \frac{\beta}{2} - \frac{\beta}{2}\right)n(2p - n + 1) + \frac{\beta}{4}n(n - 1),$$
we only need to check $A' + B' = 0$. In fact, write

$$A' = \frac{-\beta np}{2} \log 2 + \left( \frac{1}{2} \log \frac{\beta}{2} \right) n - \frac{(\log \beta)n}{4} (2p - n + 1) \beta + \frac{\beta (\log 2)n}{4} (2p - n + 1)$$

$$+ \frac{\beta}{4} n (2p - n + 1) + \frac{\beta}{4} n (n - 1)$$

(3.117)

$$= -\frac{(\log \beta)n}{4} (2p - n + 1) \beta + \frac{1}{2} (\log \beta) n + \left( -\frac{\beta n (n - 1)}{4} - \frac{n}{2} \right) \log 2 + \frac{1}{2} np \beta$$

where in the second equality we first merge all terms with log 2 together by writing

$$\left( \frac{1}{2} \log \frac{\beta}{2} \right) n = \frac{1}{2} (\log \beta) n - \frac{1}{2} (\log 2) n,$$

and then sum the last two terms in (3.117) to obtain $\frac{1}{2} np \beta$. Now it is evident that $A' = -B'$. ■
Chapter 4

Limiting Laws and Berry-Essen Bound of Statistics for Testing High Dimension Covariance Structure

4.1 Introduction

Let $X_1, \ldots, X_n$ be independent and identically distributed $p$-dimensional random vectors with mean $\mu = \mu_{p \times 1}$ and covariance matrix $\Sigma$, which forms a random matrix $X_{n,p}$. Testing the covariance matrix:

$$H_0: \Sigma = \sigma^2 I_p \quad \text{vs.} \quad H_1: \Sigma \neq \sigma^2 I_p,$$

where $I_p$ is the $p$-dimensional identity matrix and $\sigma^2$ is an unknown but finite positive constant, which is practical needs come from many fields of statistical application, Especially for Gaussian case. The traditional method based on the sample covariance $X'X$ such as the likelihood ratio test, see Anderson (2003), can not function any more when $p \to \infty$ as $n \to \infty$.

When dimension $p$ and the sample size $n$ are comparable, i.e., $n/p \to \gamma \in \mathbb{R}$,
Many methods were developed based on random theory. By assuming \( \mu = 0 \), the largest eigenvalue has been considered for testing hypothesis in (4.1) by Johnstone (2001) in the Gaussian case and by Péché (2009) in the more general case where the distribution is assumed to be sub-Gaussian tails and the ratio \( p/n \) can converge to either a positive number \( \gamma \), 0 or \( \infty \). Ledoit and Wolf (2002) first used the trace of the quadratic forms of the sample covariance as a new test statistic to test the null hypothesis under the normality assumption, and by weaken the conditions, Chen, Zhang and zhong (2010) also introduced a similar test statistic. Bai, Jiang, Yao and zheng (2009) introduced the corrections to LRT to solve this kind problem.

Without the normality assumption, still under comparable setting, Jiang (2004) first introduced the Coherence statistic \( \tilde{L}_n = \max_{1 \leq i < j \leq p} |\hat{\rho}_{ij}| \), where \( \hat{\rho}_{ij} \) is the sample correlation coefficient. Then Zhou (2007) reduced the moment conditions for the limiting law of the coherence statistic. Under some new finite moment assumption, Cai and Jiang (2011) released the comparable condition \( n/p \to \gamma > 0 \) to \( p \) can be as large as \( e^{n^{1/3}} \), and proved the limiting distribution of the coherence statistic is the extreme distribution of Type I. However, Liu, Lin and Shao (2008) already pointed out the convergence rate for the limiting distribution of the coherence statistic is as slow as \( O(\log_2 n/\log n) \), which highly influences the speed to approach the test significant level. Under the case that \( c_1n^a \leq p \leq c_2n^a \), they introduced a new test statistic in the following way: let \( X_{n,p} = (x_{ij}) \) be an \( n \times p \) random matrix where the entries \( x_{ij} \) are i.i.d. real continuous random variables with mean \( \mu \) and variance \( \sigma^2 > 0 \) which are unknown. Let \( x_1, x_2, \ldots, x_p \) denote the \( p \) columns of \( X_{n,p} \), then \( \bar{X}_{n,p} = (x_1, x_2, \ldots, x_p) \). Let \( \bar{x}_i = (1/n) \sum_{k=1}^n x_{ki} \) be the sample average of \( x_k \), and define

\[
L_n^2 = \max_{1 \leq i < j \leq p} r_{ij}^2,
\]  

(4.2)
where

\[
\begin{align*}
    r_{i,j}^2 &= (2\tilde{A}_{n,i,j}^2 + 2\tilde{B}_{n,i,j}^2)/D_{n,i,j}, \\
    \tilde{A}_{n,i,j} &= \sum_{k=1}^{[\frac{n}{2}]} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j), \\
    \tilde{B}_{n,i,j} &= \sum_{k=1+[\frac{n}{2}]}^{n} (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j), \\
    D_{n,i,j} &= \sum_{k=1}^{n} (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^{n} (x_{kj} - \bar{x}_j)^2
\end{align*}
\]

(4.3)

where \([\frac{n}{2}]\) denotes the integer part of \(\frac{n}{2}\). They proved the corresponding limiting result of it. From now on, we call this test statistic as the modified coherence statistic.

Motivated by the possible improvement of the convergence speed of the modified coherence statistic. In our paper, instead of assuming \(c_1 n^a \leq p \leq c_2 n^a\), we consider the high dimension case where \(p\) can be as large as \(c n^\beta\) for some \(0 < \beta < \frac{1}{3}\) with different moment assumption. We will show that the modified coherence \(L_n^2\) has an extreme limit distribution with a good convergence rate \(\sqrt{\frac{\log p}{n^2}}\), then we can use \(L_n^2\) to test null hypothesis \(H_0 : \sigma_{ij} = 0\) for all \(|i - j| \geq 1\) instead of (4.1). In addition, we also consider the case where the enitres of gaussian random matrix \(X_{n,p}\) are correlated. Let \(X_{n,p} = (x_{ij})\), where the \(n\) rows are i.i.d. random vectors distributed with mean \(\mu\) and covariance \(\Sigma\), where \(\mu \in \mathbb{R}^p\) is arbitrary in this section unless otherwise specified. Then define an analogous statistic:

\[
L_{n,\tau}^2 = \max_{|i - j| \geq \tau} r_{i,j}^2
\]

(4.4)

where \(r_{i,j}^2\) is same as in (4.3). We can use \(L_{n,\tau}\) to test the covariance matrix \(\Sigma\) is banded, which is

\[
H_0 : \sigma_{ij} = 0 \text{ for all } |i - j| \geq \tau
\]

(4.5)

for a given integer \(\tau \geq 1\). Under the appropriate conditions, the limiting distribution be obtained with a good convergence rate.
The paper is structured as follows. Section 2 introduces the main results and applications of the main results for testing. The proofs of the main result are given in the section 3. Some technical details and useful lemmas are deferred to the Section 4 and The Appendix.

4.2 Main Results and Applications

In this section, we consider the limiting laws and the Berry-Essen bound of the modified coherence $L^2_n$ of a random matrix with i.i.d. entries. In addition, we also investigate $L^2_{n,\tau}$ which formed from a random matrix where each row of it is drawn independently from a multivariate Gaussian distribution with banded covariance matrix. In the latter case, we also give out the limiting distribution and corresponding bound. Then apply the asymptotic results to the testing of the covariance structure in Section 4.2.3.

4.2.1 The Independent Case

The first result states the law of large numbers for $L^2_n$ which is defined in (4.2) for the case where the random entries $x_{ij}$ are bounded i.i.d..

**THEOREM 4.1** Assume $|x_{11}| \leq C$ for a finite constant $C > 0$, and $p = p(n) \to \infty$ and $\log p = o(n)$ as $n \to \infty$. Then $nL^2_n \log p \to 4$ in probability as $n \to \infty$. We now consider the case where $x_{ij}$ are i.i.d with finite exponential moments.

**THEOREM 4.2** Suppose $Ee^{t_0|x_{11}|^\alpha} < \infty$ for some $\alpha > 0$ and $t_0 > 0$. Set $\beta = \alpha/(4 + \alpha)$. Assume $p = p(n) \to \infty$ and $\log p = o(n^\beta)$ as $n \to \infty$. Then $nL^2_n \log p \to 4$ in probability as $n \to \infty$.

Comparing Theorems 4.1 and 4.2, it can be seen that a stronger moment condition gives a higher order of $p$ to make the law of large numbers for $L_n$ valid. Also, based on Theorem 4.2, if $Ee^{t|x_{11}|^\alpha} < \infty$ for any $\alpha > 0$, then $\beta \to 1$, hence the order $o(n^\beta)$ is close to $o(n)$, which is the order in Theorem 4.1.
We now consider the limiting distribution of $L_n^2$ after suitable normalization.

**Theorem 4.3** Suppose $Ee^{t|x_i|} < \infty$ for some $0 < \alpha \leq 2$ and $t_0 > 0$. Set $\beta = \alpha/(4 + \alpha)$. Assume $p = p(n) \to \infty$ and $\log p = o(n^\beta)$ as $n \to \infty$. Then $nL_n^2 - 4n\log p$ converges weakly to an extreme distribution of type I with distribution function
\[ F(y) = \exp\left(-\frac{1}{2}e^{-\frac{y}{2}}\right), \]
and if $\log n/\log p \to 0$,
\[
\sup_{y \in \mathbb{R}} \left| P(nL_n^2 - 4\log p \leq y) - \exp\left(-\frac{1}{2}e^{-\frac{y}{2}}\right) \right| \leq \sqrt{\frac{\log p^3}{n}}.
\]

**Remark 4.1** Theorem 4.3 will still hold for $\alpha > 2$, in that case, we only need to consider $\alpha$ as $2$.

### 4.2.2 The Dependent Case

As mentioned in the introduction, it is of interest to test the hypothesis in (4.5) that the covariance matrix $\Sigma$ is banded. In order to construct a test, we study the asymptotic distributions of $L_{n,\tau}$ defined in (4.4) in this section. Assuming the covariance matrix $\Sigma$ has desired banded structure under the null hypothesis and let $R = (\rho_{ij})_{p \times p}$ be the correlation matrix obtained from $\Sigma$. This case is much harder than the i.i.d. case considered in Section 4.2.1 because of the dependence. So we define, for any $0 < \delta < 1$,
\[
\Gamma_{p,\delta} = \{1 \leq i \leq p; \ |\rho_{ij}| > 1 - \delta \text{ for some } 1 \leq j \leq p \text{ with } j \neq i\}. \tag{4.6}
\]

**Theorem 4.4** Suppose, as $n \to \infty$,
(i) $p = p_n \to \infty$ with $\log p = o(n^{1/3})$;
(ii) $\tau = o(p^t)$ for any $t > 0$;
(iii) for some $\delta \in (0, 1)$, $|\Gamma_{p,\delta}| = o(p)$, which is particularly true if $\max_{1 \leq i < j \leq p < \infty} |\rho_{ij}| \leq 1 - \delta$. 


Then, under $H_0$ in (4.5), $nL^2_{n,\tau} - 4\log p$ converges weakly to an extreme distribution of type I with distribution function

$$F(y) = \exp(-\frac{1}{2}e^{-\frac{y}{2}}), \quad y \in \mathbb{R}.$$ 

**THEOREM 4.5** Suppose, as $n \to \infty$,

(i) $p = p_n \to \infty$ with $\log p = o(n^{1/3})$ and $\frac{\log n}{\log p} \to 0$;

(ii) $\tau = o(p^{t})$ for any $t > 0$;

(iii) for some $\delta \in (0, 1)$, $\frac{|G_{p,\delta}|}{p} \leq O\left(\sqrt{\frac{\log p}{n}}\right)$, which is particularly true if $\max_{1 \leq i < j \leq p} |\rho_{ij}| \leq 1 - \delta$.

Then, under $H_0$, $nL^2_{n,\tau} - 4\log p$ converges weakly to an extreme distribution of type I with distribution function,

$$\left|P(nL^2_{n,\tau} - 4\log p \leq y) - \exp(-\frac{1}{2}e^{-\frac{y}{2}})\right| \leq \sqrt{\frac{(\log p)^3}{n}}.$$ 

### 4.2.3 Testing the Covariance Structure

The limiting laws derived in the last two sections have immediate statistical applications. We only consider the Gaussian case as an example. Observe independent and identically distributed $p$-variate Gaussian variables $Y_1, \ldots, Y_n$ with mean $\mu_{p \times 1}$, covariance matrix $\Sigma_{p \times p} = (\sigma_{ij})$ and correlation matrix $R_{p \times p} = (r_{ij})$. For a given integer $\tau \geq 1$ and a given significant level $0 < \alpha < 1$, instead of testing (4.1), we wish to test the hypotheses

$$H_0 : \sigma_{i,j} = 0 \text{ for all } |i - j| \geq \tau \text{ versus } H_a : \sigma_{i,j} \neq 0 \text{ for some } |i - j| \geq \tau. \quad (4.7)$$

A case of special interest is $\tau = 1$, which corresponds to testing independence of the Gaussian random variables. The asymptotic distribution of $L_{n,\tau}$ derived in Section 4.2.2 can be used to construct a convenient test statistic for testing the hypotheses in (4.7).

Based on the asymptotic result given in Theorem 4.4 that

$$P \left( nL^2_{n,\tau} - 4\log p \leq y \right) \to \exp(-\frac{1}{2}e^{-\frac{y}{2}}), \quad (4.8)$$

$$P \left( nL^2_{n,\tau} - 4\log p \leq y \right) \to \exp(-\frac{1}{2}e^{-\frac{y}{2}}).$$
we define a test for testing the hypotheses in (4.7) by

\[ T_\tau = I \left( L_{n,\tau}^2 \geq n^{-1}(4 \log p - \log 4 - 2 \log \log(1 - \alpha)^{-1}) \right). \]  

(4.9)

That is, we reject the null hypothesis \( H_0 \) with asymptotic size \( \alpha \) whenever

\[ L_{n,\tau}^2 \geq n^{-1}(4 \log p - \log 4 - 2 \log \log(1 - \alpha)^{-1}) \]

Note that for \( \tau = 1 \), \( L_{n,\tau}^2 \) reduces to \( L_n^2 \) and the test is then based on the modified coherence \( L_n^2 \).

**THEOREM 4.6** Under the conditions of Theorem 4.3, the test \( T_1 \) defined in (4.9) has size \( \alpha \) asymptotically.

**THEOREM 4.7** Under the conditions of Theorem 4.4, the test \( T_\tau \) defined in (4.9) has size \( \alpha \) asymptotically.

These result are direct consequence of Theorem 4.3 and 4.4.

### 4.3 Proofs

In this section, we prove Theorems 4.1 - 4.5. The letter \( C \) stands for a constant and may vary from place to place throughout this section. Also, we sometimes write \( p \) for \( p_n \) if there is no confusion.

#### 4.3.1 Proofs of Theorems 4.1, 4.2 and 4.3

Define

\[ W_n^2 = \max_{1 \leq i < j \leq p} 2(A_{n,i,j}^2 + B_{n,i,j}^2) \]  

(4.10)

where

\[ A_{n,i,j} = \sum_{k=1}^{[\frac{n}{2}]} x_{k,i}x_{k,j}, \quad B_{n,i,j} = \sum_{k=1+\left[\frac{n}{2}\right]}^{n} x_{k,i}x_{k,j} \]  

(4.11)

with \( \left[\frac{n}{2}\right] \) denotes the integer part of \( n/2 \).

The following Lemmas are essential for our proof.
**Lemma 4.1** Recall the definitions of $L_n^2$ and $W_n^2$ from (4.2) and (4.10). Let $h_i = \left( \sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \right)^{\frac{1}{2}} / \sqrt{n}$. Assume that

$$b_{n,1} = \max_{1 \leq i \leq p} |h_i - 1|, \quad b_{n,2} = \min_{1 \leq i \leq p} h_i, \quad b_{n,3} = \max_{1 \leq i \leq p} |\bar{x}_i|$$

and $b_{n,4} = \max_{1 \leq i \leq p} (|x_{i1}^{(1)}|, |x_{i1}^{(2)}|)$, where $\bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{ki}$, $x_i^{(1)} = \frac{\sum_{k=1}^n x_{ki}}{n-i}$ and $x_i^{(2)} = \frac{\sum_{k=1}^n x_{ki}}{n-i-1}$. Then

$$|n^2 L_n^2 - W_n^2| \leq (B_1 + 2B_1)W_n^2 + 4(B_1B_2 + B_2)W_n| + 4B_2^2,$$

where $B_1 = (b_{n,1}^2 + 2b_{n,1})b_{n,2}, B_2 = (2 + n)b_{n,2}^2(b_{n,3} + b_{n,4})$.

**Lemma 4.2** Let $\{x_{ij}; i \geq 1, j \geq 1\}$ be i.i.d. random variables with $E x_{11} = 0$ and $E x_{11}^2 = 1$. Then, $b_{n,2} \to 1$ in probability as $n \to \infty$, and $\{\sqrt{n/ \log p} b_{n,3}\}$, $\{\sqrt{n/ \log p} b_{n,4}\}$, $\{\sqrt{n/ \log p} B_1\}$ and $\{B_2/ \log p\}$ are tight provided one of the following conditions holds:

(i) $|x_{11}| \leq C$ for some constant $C > 0$, $p_n \to \infty$ and $\log p_n = o(n)$ as $n \to \infty$;

(ii) $E e^{t \alpha |x_{11}|^\alpha} < \infty$ for any $\alpha > 0$ and $t > 0$, and $p_n \to \infty$ and $\log p_n = o(n^\beta)$ as $n \to \infty$, where $\beta = \alpha/(4 + \alpha)$.

**Lemma 4.3** Let $\xi_1, \ldots, \xi_n$ be i.i.d. random variables with $E \xi_1 = 0$, $E \xi_1^2 = 1$ and $E e^{t_0 \alpha |\xi_1|}$ < $\infty$ for some $t_0 > 0$ and $0 < \alpha_1 \leq 1$. Put $S_n = \sum_{i=1}^n \xi_i$, $T_n = \sum_{i=1}^n \xi_i$ and $\beta = \alpha_1/(2 + \alpha_1)$. Then, for any $\{p_n; n \geq 1\}$ with $0 < p_n \to \infty$ and $\log p_n = o(n^\beta)$ and $\{y_n; n \geq 1\}$ with $y_n \to y > 0$,

$$|P\left(\frac{(S_n^2 + T_n^2)}{n \log p_n} \geq y_n^2 \right) - \exp\left(-\frac{y_n^2}{2} \log p_n\right)| \sim O\left(\frac{(\log p_n)^\beta}{n \log p_n} y_n^2 \right).$$

**Lemma 4.4** Suppose $\{x_{ij}; i \geq 1, j \geq 1\}$ are i.i.d. random variables with $E x_{11} = 0$, $E (x_{11})^2 = 1$ and $E e^{t_0 |x_{11}|}$ < $\infty$ for some $t_0 > 0$ and $0 < \alpha \leq 2$. Assume $p = p(n) \to \infty$ and $\log p = o(n^\beta)$ as $n \to \infty$, where $\beta = \alpha/(4 + \alpha)$. Let $A_{n,ij}$ and $B_{n,ij}$ follow the definition as in (4.11), then for any $\epsilon > 0$ and a sequence of positive numbers $\{t_n\}$ with limit $t > 0$,

$$\Psi_n := P\left(\frac{2A_{n,1.2}^2 + 2B_{n,1.2}^2}{n \log p} > t_n^2, \frac{2A_{n,1.3}^2 + 2B_{n,1.3}^2}{n \log p} > t_n^2\right) = O\left(\frac{1}{p^{t^2 - \epsilon}}\right)$$

as $n \to \infty$. 
The proofs of them will be given in the section 4.4

**Proposition 4.1** Suppose \( \{x_{ij}; i \geq 1, j \geq 1\} \) be i.i.d. random variables with \( |x_{11}| \leq C \) for a finite constant \( C > 0 \), \( E x_{11} = 0 \) and \( E(x_{11}^2) = 1 \). Assume \( p = p(n) \to \infty \) and \( \log p = o(n) \) as \( n \to \infty \). \( W_n^2 \) is defined as in (4.10). Then

\[
\frac{W_n^2}{n \log p} \to 4
\]

in probability as \( n \to \infty \).

**Proof.** According to the definition of \( H_n \) as in lemma 4.14, we know that \( W_n^2 \geq H_n^2 \). By the lemma 4.14, we conclude that, for any \( \epsilon > 0 \),

\[
\lim_{n \to \infty} P\left( \frac{W_n^2}{n \log p} \leq 4 - \epsilon \right) = 0.
\]

We only need to prove

\[
\lim_{n \to \infty} P\left( \frac{W_n^2}{n \log p} \geq 4 + 4\epsilon \right) = 0 \quad (4.12)
\]

for any \( \epsilon > 0 \). By the definition of \( W_n^2 \),

\[
P(W_n^2 \geq (4 + 4\epsilon)n \log p) \leq \left( \frac{p}{2} \right) \cdot P\left( A_{n,1,2}^2 + B_{n,1,2}^2 \geq (2 + 2\epsilon)n \log p \right) \quad (4.13)
\]

for any \( \epsilon > 0 \). First assume \( n \) is even, let \( \xi_k = (x_{k1}x_{k2}, x_{(k+\frac{n}{2})1}x_{(k+\frac{n}{2})2})' \), then it is a i.i.d. random vectors with \( E\xi_k = 0 \), \( \|\xi_k\| \leq 2C^2 \) and \( \lambda_1 = 1 = \lambda_2 \) are the eigenvalues of \( \text{Var}(\xi_k) \). By the lemma 4.9,

\[
P\left( A_{n,1,2}^2 + B_{n,1,2}^2 \geq (2 + 2\epsilon)n \log p \right)
\]

\[
= P\left( \frac{\|\sum_{k=1}^{\frac{n}{2}} \xi_k\|^2}{\sqrt{n/2}} \geq \sqrt{(4 + 4\epsilon) \log p} \right)
\]

\[
\leq \sqrt{\frac{\pi}{2}} (4 + 4\epsilon) \log p \exp\left( - \frac{(4 + 4\epsilon) \log p}{2} (1 + \frac{a}{3})^{-1} \right),
\]

where \( a = \frac{2C^2 \sqrt{(4 + 4\epsilon) \log p}}{\sqrt{n}} \). Since \( \log p = o(n) \), so when \( n \) is sufficiently large, then \( a \leq \epsilon/2 \),

\[
P\left( A_{n,1,2}^2 + B_{n,1,2}^2 \geq (2 + 2\epsilon)n \log p \right) \leq C \frac{1}{p^{2+\epsilon}}. \quad (4.14)
\]
Using (4.18) and (4.14), then we get

\[ P(W_n^2 \geq (4 + 4\epsilon)n \log p) \leq \frac{1}{pe} \rightarrow 0 \quad (4.15) \]

as \( n \rightarrow \infty \). So we can conclude that

\[ \frac{W_n^2}{n \log p} \rightarrow 4 \]

as \( n \rightarrow \infty \). If \( n \) is odd,

\[ W_n^2 \leq W_{n-1}^2 + 4 \max_{1 \leq i < j \leq p} |x_{ni}x_{nj}B_{n-1,i,j}| + 2 \max_{1 \leq i < j \leq p} (x_{ni}x_{nj})^2, \quad (4.16) \]

where \( |B_{n-1,i,j}| \leq H_{n-1} \) which is defined in Lemma 4.14 and it is independent of \( x_{ni}x_{nj} \). As we know \( |x_{ij}| \leq C \), together with Lemma 4.14, we have

\[ P(W_n^2 \geq (4 + 4\epsilon)n \log p) \leq P(W_{n-1}^2 \geq (4 + 2\epsilon)(n - 1) \log p) + P(H_n \geq \frac{en \log p}{4}) \]

\[ \leq C \frac{1}{pe} \rightarrow 0, \]

when \( n \) is large enough. Thus (4.12) holds. Now, we finish the proof. \( \blacksquare \)

**Proposition 4.2** Suppose \( \{x_{ij}; 1 \geq i, j \geq 1\} \) are i.i.d. random variables with \( Ex_{11} = 0 \), \( E(x_{11}^2) = 1 \) and \( Ee^{t_0|x_{11}|^\alpha} < \infty \) for some \( t_0 > 0 \) and \( \alpha > 0 \). Assume \( p = p(n) \rightarrow \infty \) and \( \log p = o(n^\beta) \) as \( n \rightarrow \infty \), where \( \beta = \alpha/(4 + \alpha) \). Let \( W_n^2 \) be as in (4.10), then

\[ \frac{W_n^2}{n \log p} \rightarrow 4 \]

in probability as \( n \rightarrow \infty \).

**Proof of Proposition 4.2.** Since \( W_n > H_n \), by lemma 4.15, that, for any \( \epsilon > 0 \),

\[ P\left( \frac{W_n^2}{n \log p} \leq 4 - \epsilon \right) = 0. \]
We only need to prove

\[ \lim_{n \to \infty} P\left( \frac{W_n^2}{n \log p} \geq 4 + 4\epsilon \right) = 0 \]  \tag{4.17}

for any \( \epsilon > 0 \). First assume \( n \) is even, \( P(W_n^2 \geq (4 + 4\epsilon) n \log p) \leq \left( \frac{p}{2} \right) \cdot P(A_{n,1,2}^2 + B_{n,1,2}^2 \geq (2 + 2\epsilon) n \log p) \)  \tag{4.18}

for any \( \epsilon > 0 \). Now we will prove that

\[ P\left( A_{n,1,2}^2 + B_{n,1,2}^2 \geq (2 + 2\epsilon) n \log p \right) = O(p^{-2-\epsilon}). \]

**Step 1.** Set \( r_n = n^{1-\beta/2} \), \( \mu_n = E x_k x_k I(|x_k x_k| \leq r_n) \). Define

\[ y_k = x_k x_k I(|x_k x_k| \leq r_n) - \mu_n, \quad z_k = x_k x_k I(|x_k x_k| > r_n) \]

for all \( i \geq 1 \) and \( j \geq 1 \). Let \( \sigma_n^2 = var(y_k) \). We first claim that there exists a constant \( C > 0 \) such that

\[ \max \left\{ |\mu_n|, |\sigma_n^2 - 1|, P(|x_k x_k| > n^\gamma) \right\} \leq C e^{-n^\beta/C} \]  \tag{4.19}

for all \( n \geq 1 \). In fact, since \( E x_k x_k = 0 \) and \( \alpha \gamma/2 = \beta \),

\[ |\mu_n| \leq E \left( |x_k x_k| e^{t_0|x_k x_k|^\frac{\alpha}{2}} \right) e^{-t_0 n^{\beta/2}} \leq E \left( |x_k x_k| e^{t_0(x_k^2 + x_k^2)} \right) e^{-t_0 n^{\beta/2}} \]  \tag{4.20}

for all \( n \geq 1 \). Note that \( |\sigma_n^2 - 1| = \mu_n^2 + E(x_k x_k)^2 I(|x_k x_k| > r_n) \), by the same argument as in (4.20), we know both \( |\sigma_n^2 - 1| \) and \( P(|x_k x_k| > r_n) \) are bounded by \( C e^{-n^\beta/C} \) for some \( C > 0 \). Then (4.19) follows.

Let \( \xi_k = (x_k x_k, x_k x_k) \), \( \eta_k = (y_k, y_k) \) and \( \theta_k = (z_k + \mu_n, z_k + \gamma + \mu_n) \), where \( \xi_k = \eta_k + \theta_k \). For any \( \epsilon > 0 \), It is easy to see

\[
\begin{align*}
P(A_{n,1,2}^2 + B_{n,1,2}^2 \geq (2 + 2\epsilon)n \log p) &= P\left( \left\| \sum_{k=1}^{n/2} \xi_k \right\| \geq \sqrt{(2 + 2\epsilon)n \log p} \right) \\
&\leq P\left( \left\| \sum_{k=1}^{n/2} \eta_k \right\| \geq \sqrt{(2 + \epsilon)n \log p} \right) \\
&+ P\left( \left\| \sum_{k=1}^{n/2} \theta_k \right\| \geq \sqrt{\epsilon n \log p} \right). \tag{4.21}
\end{align*}
\]
Step 2: Now we need to consider the two equations (4.21) and (4.21) when n is even. Since $|y_k| \leq 2r_n, Ey_k = 0$ and $Ey_k^2 = \sigma_n^2$ for all $k \geq 1$. We use the Bernstein inequality in Lemma 4.9 on $\eta_k$, while $\lambda_1 = \sigma_n^2 = \lambda_2, H = 2r_n^2, r = \sqrt{(4 + 2\epsilon) \log p}$ and $a = \frac{4H}{\sqrt{n}H_1}$, then

$$P(\|\frac{1}{2} \sum_{k=1}^{n/2} \eta_k \| \geq \sqrt{(4 + 2\epsilon) \log p \sqrt{n/2}})$$

$$\leq \sqrt{\frac{\pi}{2\sigma_n^2}} (4 + 2\epsilon) \log p \exp \left(-\frac{(4 + 2\epsilon) \log p}{2\sigma_n^2}(1 + \frac{a}{3})^{-1}\right).$$

When n is large enough, $a \leq \epsilon/8, \sigma_n \leq 1 + \epsilon/8$, then

$$P(\|\frac{1}{2} \sum_{k=1}^{n/2} \eta_k \| \geq \sqrt{(2 + \epsilon)n \log p}) \leq CP^{-2-\epsilon}. \quad (4.22)$$

Recalling the definition of $z_k, \mu_n$ and $\theta_k$, we have

$$P(\|\sum_{k=1}^{n/2} \theta_k \| \geq \sqrt{\epsilon n \log p})$$

$$\leq P(\|\sum_{k=1}^{n/2} (z_k, z_{k+n/2}) \| > \sqrt{\epsilon n \log p / 2}) + P(|u_n| > \sqrt{\epsilon \log p / n})$$

$$\leq P(\max_{1 \leq k \leq n} \|x_k x_{k+1} \| \geq r_n) + P(|u_n| > \sqrt{\epsilon \log p / n}). \quad (4.23)$$

Then, by (4.19), when n is large enough,

$$P(\|\sum_{k=1}^{n/2} \theta_k \| \geq \sqrt{\epsilon n \log p}) \leq Ce^{-n^\beta / C}. \quad (4.24)$$

Combining (4.22) and (4.24), $\log p = o(n^\beta), \quad P(W^2_n \geq (4 + 4\epsilon) n \log p) \leq \frac{1}{p^\epsilon} \rightarrow 0, \quad (4.25)$

when n is even and large enough.
Step 3: Now we consider the case $n$ is odd. Since $E \exp(t_0 x_{ni}^\alpha) < \infty$, by Lemma 4.15, (4.16) and using (4.25), we get

$$P(W_n^2 \geq (4 + 4\epsilon)n \log p) \leq P(W_{n-1}^2 \geq (4 + 2\epsilon)(n - 1) \log p) + P(H_n > 4\sqrt{n \log p}) + 2P(16 \max_{1 \leq i < j \leq p} |x_{ni}x_{nj}| \geq \epsilon \sqrt{n \log p}) \leq C \left( \frac{1}{p^\epsilon} + \frac{1}{p^2} + \frac{1}{p^{(\log p)^{1/2}} - 2} \right) \to 0$$

as $n \to \infty$. While the second inequality is due to $n^{\alpha/4} > \log p$ and

$$P(\max_{1 \leq i < j \leq p} |x_{ni}x_{nj}| \geq \frac{\epsilon \sqrt{n \log p}}{16}) \leq p^2 \exp(-\frac{\epsilon}{16} n^{2/3} (\log p)^{2/3}) \leq \exp(-\frac{\epsilon}{16} \log p(\log p)^{2/3}).$$

Now we proved (4.17), then the Proposition is proved. 

Proof of Theorem 4.1. First, for constants $\mu_i \in \mathbb{R}$ and $\sigma_i > 0$, $i = 1, 2, \cdots, p$, it is easy to see that matrix $X_{n,p} = (x_{ij})_{n \times p} = (x_1, x_2, \cdots, x_p)$ and $(\sigma_1 x_1 + \mu_1 e, \sigma_2 x_2 + \mu_2 e, \cdots, \sigma_p x_p + \mu_p e)$ generate the same $L_n^2$ as in (4.2), where $e = (1, \cdots, 1)' \in \mathbb{R}^n$. Thus, w.l.o.g., we prove the theorem next by assuming that $\{x_{ij}; 1 \leq i \leq n, 1 \leq j \leq p\}$ are i.i.d. random variables with mean zero and variance 1.

By Proposition 4.1, under condition $\log p = o(n)$,

$$\frac{W_n^2}{n \log p} \to 4 \quad (4.26)$$

in probability as $n \to \infty$. Thus, to prove the theorem, it is enough to show

$$\frac{n^2 L_n - W_n^2}{n \log p} \to 0 \quad (4.27)$$

in probability as $n \to \infty$. From Lemma 4.1, we get that

$$|n^2 L_n^2 - W_n^2| \leq (B_1^2 + 2B_1)W_n^2 + 2(2B_1 B_2 + 2B_2)W_n + 4B_2^2.$$
See the Lemma 4.2 under the condition (i) and Proposition 4.1,
\[ \frac{W_n}{\sqrt{n \log p}} \to 2, \quad \sqrt{\frac{n}{\log p}} B_1 \text{ and } \frac{B_2}{\log p} \text{ are tight.} \]
Recall the definition of tight, we can conclude (4.27) is true. ■

**Proof of Theorem 4.2.** In the proof of Theorem 4.1, replace “Proposition 4.1” with “Proposition 4.2” and “(i) of Lemma 4.2” with “(ii) of Lemma 4.2”, keep all other statements the same, we then get the desired result. ■

**Proposition 4.3** Let \( \{x_{ij}; i \geq 1, j \geq 1\} \) be i.i.d. random variables with \( E x_{11} = 0, E(x_{11}^2) = 1 \) and \( E e^{t|x_{11}|^p} < \infty \) for some \( 0 < \alpha \leq 2 \) and \( t_0 > 0 \). Set \( \beta = \min (\alpha/(4 + \alpha), \alpha/(6 - \alpha)) \). Assume \( p = p(n) \to \infty \) and \( \log p = o(n^\beta) \) as \( n \to \infty \). Then
\[ P \left( \frac{W_n^2 - 4n \log p}{n} \leq z \right) \to \exp \left( -\frac{1}{2} e^{-\frac{z}{2}} \right) \]
as \( n \to \infty \). Furthermore,
\[ \sup_{z \in \mathbb{R}} \left| P \left( \frac{W_n^2 - \alpha_n}{n} \leq z \right) - \exp \left( -\frac{1}{2} e^{-\frac{z}{2}} \right) \right| \leq C \left( \frac{1}{p^{1-\epsilon}} + \sqrt{\log p^3 / n} \right). \]

**Proof.** Let \( \alpha_n = 4n \log p \), first we show that
\[ P \left( \frac{W_n^2 - 4n \log p}{n} \leq z \right) = P \left( \max_{1 \leq i<j \leq p} y_{ij} \leq \alpha_n + nz \right) \to \exp \left( -\frac{1}{2} e^{-\frac{z}{2}} \right), \]
where \( y_{ij} = 2A_{n,i,j}^2 + 2B_{n,i,j}^2 \). We will use Lemma 4.10 to prove it. Take \( I = \{(i,j); 1 \leq i < j \leq p\}, \) for \( u = (i,j) \in I \), set \( X_u = y_{ij} \) and \( B_u = \{(k,l) \in I; \text{ one of } k \text{ and } l = i \text{ or } j, \text{ but } (k,l) \neq u\}. \) Let \( a_n = \alpha_n + nz \) and \( C_{ij} = \{y_{ij} > a_n\} \). Since \( \{y_{ij}; (i,j) \in I\} \) are identically distributed, by Lemma 4.10,
\[ |P(W_n^2 \leq a_n) - e^{-\lambda_n}| \leq b_{1,n} + b_{2,n}, \]
where
\[ \lambda_n = \frac{p(p-1)}{2} P(C_{12}), \quad b_{1,n} \leq 2p^3 P(C_{12})^2 \text{ and } b_{2,n} \leq 2p^3 P(C_{12}C_{13}). \]
We first calculate $\lambda_n$, write

$$\lambda_n = \frac{p^2 - p}{2} P \left( y_{12} > \alpha_n + n z \right), \quad (4.32)$$

where $y_{12} = 2(\sum_{i=1}^{\lceil \frac{n}{2} \rceil} \xi_i)^2 + 2(\sum_{i=\lceil \frac{n}{2} \rceil + 1}^{n} \xi_i)^2$, and $\{\xi_i; 1 \leq i \leq n\}$ are i.i.d. random variables with the same distribution as that of $x_{11}x_{12}$. In particular, $E\xi_1 = 0$ and $E\xi_1^2 = 1$. Note $\alpha_1 := \alpha / 2 \leq 1$. We then have

$$|x_{11}x_{12}|^{\alpha_1} \leq \left( \frac{x_{11}^2 + x_{12}^2}{2} \right)^{\alpha_1} \leq \frac{1}{2^{\alpha_1}} \left( |x_{11}|^{\alpha} + |x_{12}|^{\alpha} \right).$$

Hence, by independence,

$$E e^{t_0 |\xi_1|^{\alpha_1}} = E e^{t_0 |x_{11}x_{12}|^{\alpha_1}} < \infty.$$  

Let $z_n = \sqrt{\left( \frac{\alpha_n}{n} + z \right) / \log p}$, then $z_n \to 2$ as $n \to \infty$. Now we write

$$P \left( \frac{y_{12}}{n} > \frac{\alpha_n}{n} + z \right) = P \left( \frac{2(\sum_{i=1}^{\lceil \frac{n}{2} \rceil} \xi_i)^2 + 2(\sum_{i=\lceil \frac{n}{2} \rceil + 1}^{n} \xi_i)^2}{n \log p} > z_n^2 \right),$$

then by Lemma 4.3,

$$|P \left( \frac{y_{12}}{n} > \frac{\alpha_n}{n} + z \right) - \exp\left( -\frac{z_n^2}{2 \log p} \right)| \leq C \sqrt{\frac{(\log p_n)^3}{n} p^{-2}}$$

as $n \to \infty$. Combining with (4.32) imply that

$$|\lambda_n - \frac{1}{2} e^{-\frac{z}{2}}| \leq C \sqrt{\frac{(\log p_n)^3}{n} + \frac{1}{p}} \quad (4.33)$$

as $n \to \infty$.

Recall (4.30) and (4.31), to complete the proof, we have to verify that $b_{1,n} \to 0$ and $b_{2,n} \to 0$ as $n \to \infty$. By (4.31), (4.32) and (4.33),

$$b_{1,n} \leq 2p^3 P(C_{12})^2 \leq \frac{8p^3 \lambda_n^2}{(p^2 - p)^2} = O \left( \frac{1}{p} \right)$$
as \( n \to \infty \). Also by (4.31),

\[
b_{2,n} \leq 2p^3 P \left( y_{12} > \alpha_n + nz, \ y_{13} > \alpha_n + nz \right)
\]

\[
= 2p^3 E \left\{ P^1 \left( \frac{2A_{n,1,2}^2 + 2B_{n,1,2}^2}{n \log P} > \frac{y_n^2}{2}, \frac{2A_{n,1,3}^2 + 2B_{n,1,3}^2}{n \log P} > \frac{y_n^2}{2} \right) \right\}
\]

where \( P^1 \) stands for the conditional probability given \( \{x_{k,1}; 1 \leq k \leq n\} \), and

\[
y_n := \sqrt{\alpha_n + nz}/\sqrt{n \log p} \to 2.
\]

By Lemma 4.4, the above expectation is equal to \( O(p^{-4}) \) as \( n \to \infty \) for any \( \epsilon > 0 \). Now choose \( \epsilon \in (0,1) \), then \( b_{2,n} = O(p^{-1}) \to 0 \) as \( n \to \infty \),

\[
|P \left( \frac{W_n^2 - \alpha_n}{n} \leq z \right) - \exp\left( -\frac{1}{2} e^{-\frac{z}{2}} \right)|
\]

\[
\leq |P \left( \frac{W_n^2 - \alpha_n}{n} \leq z \right) - e^{-\lambda_n}| + |e^{-\lambda_n} - \exp\left( -\frac{1}{2} e^{-\frac{z}{2}} \right)|
\]

\[
\leq C \left( \frac{1}{p^{1-\epsilon}} + \sqrt{\frac{(\log p_n)^3}{n}} \right).
\]

The proof is completed. \( \blacksquare \)

**Proof of Theorem 4.3.** Without loss of generality, assume \( \mu = 0 \) and \( \sigma = 1 \).

By the equation (4.27) in Theorem 4.1, we know that

\[
\left| \frac{n^2 L_n^2 - W_n^2}{n} \right| \to 0 \quad (4.34)
\]

in probability as \( n \to \infty \). When \( n \to \infty \), using Proposition 4.3, we get

\[
P \left( \frac{n^2 L_n^2 - \alpha_n}{n} \leq z \right) \to \exp\left\{ -\frac{e^{-\frac{z}{2}}}{2} \right\}.
\]

Now we prove the convergence rate. By the equations (4.71) and (4.72) in the
proof of Lemma 4.2 and Lemma 4.14, we can find a suitable $K$ and get

\[
P \left( \left| \frac{n^2 L_n^2 - W_n^2}{n} \right| \geq K \sqrt{\frac{(\log p)^3}{n}} \right)
\]

\[
\leq P \left( \frac{(B_1^2 + 2B_1)W_n^2 + 4(B_1B_2 + B_2)W_n + 4B_2^2}{n \log p} \geq K \sqrt{\frac{\log p}{n}} \right)
\]

\[
\leq P \left( \sqrt{\frac{n}{\log p}} B_1 \geq \frac{K - 1}{24} \right) + P \left( \frac{B_2}{\log p} \geq \frac{K - 1}{20} \right) + C p^{-1+\epsilon}
\]

\[
\leq C \frac{2}{p^{1-\epsilon}},
\]

when $n$ is large enough. It is easy to get that

\[
P \left( \frac{n^2 L_n^2 - \alpha_n}{n} \leq z \right)
\]

\[
\leq P \left( \frac{n^2 L_n^2 - \alpha_n}{n} \leq z, \left| \frac{n^2 L_n^2 - W_n^2}{n} \right| \leq K \sqrt{\frac{(\log p)^3}{n}} \right)
\]

\[
+ P \left( \left| \frac{n^2 L_n^2 - W_n^2}{n} \right| \geq K \sqrt{\frac{(\log p)^3}{n}} \right)
\]

\[
\leq P \left( \frac{W_n^2 - \alpha_n}{n} \leq z + K \sqrt{\frac{(\log p)^3}{n}} \right) + C \frac{2}{p^{1-\epsilon}},
\]

and by $P(AB) \geq P(A) - P(B^c)$,

\[
P \left( \frac{n^2 L_n^2 - \alpha_n}{n} \leq z \right)
\]

\[
\geq P \left( \frac{n^2 L_n^2 - \alpha_n}{n} \leq z, \left| \frac{n^2 L_n^2 - W_n^2}{n} \right| \leq K \sqrt{\frac{(\log p)^3}{n}} \right)
\]

\[
\geq P \left( \frac{W_n^2 - \alpha_n}{n} \leq z - K \sqrt{\frac{(\log p)^3}{n}} \right) - C \frac{2}{p^{1-\epsilon}}.
\]
Let $z_n = z \pm K \sqrt{\frac{(\log p)^3}{n}} \rightarrow z$ as $n \rightarrow \infty$, by the mean value theorem,

$$\left| P\left( \frac{W_n^2 - \alpha_n}{n} \leq z_n \right) - e^{-\frac{1}{2}e^{-z/2}} \right|$$

$$\leq \left| P\left( \frac{W_n^2 - \alpha_n}{n} \leq z_n \right) - e^{-\frac{1}{2}e^{-\frac{z_n}{2}}} \right| + \left| e^{-\frac{1}{2}e^{-z_n/2}} - e^{-\frac{1}{2}e^{-z/2}} \right|$$

$$\leq C\left( \frac{1}{p^{1-\epsilon}} + \sqrt{\frac{(\log p)^3}{n}} \right) + C|z_n - z|$$

$$\leq C\left( \frac{1}{p^{1-\epsilon}} + \sqrt{\frac{(\log p)^3}{n}} \right).$$

So combining the above equation and (4.35), we conclude that

$$\left| P\left( \frac{n^2L_n^2 - \alpha_n}{n} \leq z \right) - e^{-\frac{1}{2}e^{-z/2}} \right| \leq C\left( \frac{1}{p^{1-\epsilon}} + \sqrt{\frac{(\log p)^3}{n}} \right).$$

So it is to see that $\sup_{y \in R} \left| P(nL_n^2 - 4\log p \leq y) - \exp\left(-\frac{1}{2}e^{-\frac{2}{2}}\right) \right| \leq \sqrt{\frac{\log p}{n}}$ is correct when $\log n / \log p \rightarrow 0$. ■

### 4.3.2 Proof of Theorem 4.4 and 4.5

We begin to prove the Theorem 4.4 by stating four technical lemmas which are proved in the section 4.4.

Let $\{(u_{k1}, u_{k2}, u_{k3}, u_{k4})^T; 1 \leq i \leq n\}$ be a sequence of i.i.d. random vectors with distribution $N_4(0, \Sigma_4)$. Define,

$$Z_{12} = 2\left( \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} u_{k1}u_{k2} \right)^2 + 2\left( \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} u_{k1}u_{k2} \right)^2;$$

$$Z_{34} = 2\left( \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor} u_{k3}u_{k4} \right)^2 + 2\left( \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} u_{k3}u_{k4} \right)^2. \quad (4.35)$$
**Lemma 4.5** Let \( \{(u_{k1}, u_{k2}, u_{k3}, u_{k4})^T; 1 \leq i \leq n\} \) be a sequence of i.i.d. random vectors with distribution \( N_4(0, \Sigma_4) \) where

\[
\Sigma_4 = \begin{pmatrix}
1 & 0 & r_1 & 0 \\
0 & 1 & 0 & 0 \\
r_1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}, \quad |r| \leq 1.
\]

Set \( a_n = 4n \log p + ny \) for \( n \geq e^e \) and \( y \in \mathbb{R} \). Suppose \( n \to \infty, p \to \infty \) with \( \log p = o(n^{1/3}) \). Then,

\[
\sup_{|r| \leq 1} P\left( Z_{12} > a_n, Z_{34} > a_n \right) = O\left( \frac{1}{p^{1-\epsilon}} \right)
\]

(4.36) for any \( \epsilon > 0 \), where \( Z_{12}, Z_{34} \) are defined in (4.94).

**Lemma 4.6** Let \( \{(u_{k1}, u_{k2}, u_{k3}, u_{k4})^T; 1 \leq i \leq n\} \) be a sequence of i.i.d. random vectors with distribution \( N_4(0, \Sigma_4) \) where

\[
\Sigma_4 = \begin{pmatrix}
1 & 0 & r_1 & 0 \\
0 & 1 & r_2 & 0 \\
r_1 & r_2 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}, \quad |r_1| \leq 1, |r_2| \leq 1.
\]

Set \( a_n = 4n \log p + ny \) for \( n \geq e^e \) and \( y \in \mathbb{R} \). Suppose \( n \to \infty, p \to \infty \) with \( \log p = o(n^{1/3}) \). Then, as \( n \to \infty \),

\[
\sup_{|r_1|, |r_2| \leq 1} P\left( Z_{12} > a_n, Z_{34} > a_n \right) = O\left( p^{-\frac{\epsilon}{3}} \right)
\]

for any \( \epsilon > 0 \), where \( Z_{12}, Z_{34} \) are defined in (4.94).

**Lemma 4.7** Let \( \{(u_{k1}, u_{k2}, u_{k3}, u_{k4})^T; 1 \leq i \leq n\} \) be a sequence of i.i.d. random vectors with distribution \( N_4(0, \Sigma_4) \) where

\[
\Sigma_4 = \begin{pmatrix}
1 & 0 & r_1 & 0 \\
0 & 1 & 0 & r_2 \\
r_1 & 0 & 1 & 0 \\
0 & r_2 & 0 & 1
\end{pmatrix}, \quad |r_1| \leq 1, |r_2| \leq 1.
\]
Set \( a_n = 4n \log p + ny \) for \( n \geq e^e \) and \( y \in \mathbb{R} \). Suppose \( n \to \infty, p \to \infty \) with \( \log p = o(n^1/3) \). Then, for any \( \delta \in (0, 1) \), there exists \( \epsilon_0 = \epsilon(\delta) > 0 \) such that

\[
\sup_{|r_1|, |r_2| \leq 1 - \delta} P \left( Z_{12} > a_n, Z_{34} > a_n \right) = O \left( p^{-2-\epsilon_0} \right) \tag{4.37}
\]
as \( n \to \infty \), where \( Z_{12}, Z_{34} \) are defined in (4.94).

Recall notation \( \tau, \Sigma = (\sigma_{ij})_{p \times p} \) and \( X_n = (x_{ij})_{n \times p} \sim N_p(\mu, \Sigma) \) from (4.5).

**Proposition 4.4** Assume \( \mu = 0 \) and \( \sigma_{ii} = 1 \) for all \( 1 \leq i \leq p \), define \( V_n = V_{n,\tau} = \max_{1 \leq i < j \leq p, j - i \geq \tau} 2(A_{n,i,j}^2 + B_{n,i,j}^2) \). \( \tag{4.38} \)

Suppose \( n \to \infty, p = p_n \to \infty \) with \( \log p = o(n^1/3) \), \( \tau = o(p^t) \) for any \( t > 0 \), and for some \( \delta \in (0, 1) \), \( |\Gamma_{p,\delta}| = o(p) \) as \( n \to \infty \). Then, under \( H_0 \) in (4.5),

\[
P \left( \frac{V_n - \alpha_n}{n} \leq y \right) \to \exp \left( -\frac{1}{2} e^{-\frac{y}{2}} \right)
\]
as \( n \to \infty \) for any \( y \in \mathbb{R} \), where \( \alpha_n = 4n \log p \).

**Proof.** Set \( a_n = \alpha_n + ny \),

\[
\Lambda_p = \left\{ (i, j); 1 \leq i < j \leq p, j - i \geq \tau, \max_{1 \leq k \neq i \leq p} \{|\rho_{ik}|\} \leq 1 - \delta, \max_{1 \leq k \neq j \leq p} \{|\rho_{jk}|\} \leq 1 - \delta \right\} \text{ and } V_n' = \max_{(i,j) \in \Lambda_p} 2(A_{n,ij}^2 + B_{n,ij}^2). \tag{4.39}
\]

**Step 1.** We claim that, to prove the proposition, it suffices to show

\[
\lim_{n \to \infty} P \left( V_n' \leq a_n \right) = \exp \left( -\frac{1}{2} e^{-\frac{y}{2}} \right) \tag{4.40}
\]
for any \( y \in \mathbb{R} \). Notice \( \{x_{ki}, x_{kj}; 1 \leq k \leq n\} \) are \( 2n \) i.i.d. standard normals, if \( |j - i| \geq \tau \). Then

\[
P \left( V_n > a_n \right) \leq P \left( V_n' > a_n \right) + \sum_{k=1}^{[n/2]} P \left( 2 \left( \sum_{k=1}^{[n/2]} x_{k\tau+1}^2 \right)^2 + 2 \left( \sum_{k=[n/2]+1}^{n} x_{k\tau+1}^2 \right)^2 > a_n \right),
\]
where the sum runs over all pair \((i, j)\) such that \(1 \leq i < j \leq p\) and one of \(i\) and \(j\) is in \(\Gamma_{p,\delta}\). Note that \(|x_{11}x_{1\tau+1}| \leq (x_{11}^2 + x_{1\tau+1}^2)/2\), it follows that \(Ee^{|x_{11}x_{1\tau+1}|/2} < \infty\) by independence and \(E\exp(N(0, 1)^2/4) < \infty\). Since \(\{x_{k1}, x_{k\tau+1}; 1 \leq k \leq n\}\) are i.i.d. with mean zero and variance one, and \(y_n^2 := a_n/n \log p \to 4\) as \(n \to \infty\), taking \(\alpha_1 = 1\) in Lemma 4.3, we get

\[
P\left(\frac{2(\sum_{k=1}^{\lceil n/2 \rceil} x_{k1}x_{k\tau+1})^2 + 2(\sum_{k=\lceil n/2 \rceil+1}^{n} x_{k1}x_{k\tau+1})^2}{n \log p} > \frac{a_n}{n \log p}\right) 
\sim \exp -\frac{y_n^2}{2} \log p \sim \frac{e^{-y}}{p^2}
\]

as \(n \to \infty\). Moreover, note that the total number of such pairs is no more than \(2p|\Gamma_{p,\delta}|\). Therefore,

\[
P(V'_n > a_n) \leq P(V_n > a_n) 
\leq P(V'_n > a_n) + 2p|\Gamma_{p,\delta}| \cdot P\left(|2(A_{n,1,\tau+1} + B_{n,1,\tau+1}^2)| > a_n\right) 
\leq P(V'_n > a_n) + o(p^2) \cdot O\left(\frac{1}{p^2}\right)
\]

by the assumption on \(\Gamma_{p,\delta}\) and (4.41). Thus, this gives (4.40).

**Step 2.** We now apply Lemma 4.10 to prove (4.40). Take \(I = \Lambda_p\), then for \(d = (i, j) \in I\), set \(Z_d = Z_{ij} = |2A_{n,i,j}^2 + 2B_{n,i,j}^2|\), then define

\[
B_{i,j} = \{(k, l) \in \Lambda_p; |s - t| < \tau \text{ for some } s \in \{k, l\} \text{ and some } t \in \{i, j\}, \\
\text{but } (k, l) \neq (i, j)\}, a_n = 4n \log p + ny \text{ and } A_{ij} = \{|Z_{ij}| > a_n\}.
\]

It is easy to see that \(|B_{i,j}| \leq 2 \cdot (2\tau + 2\tau)p = 8\tau p\) and that \(Z_{ij}\) are independent of \(\{Z_{kl}; (k, l) \in \Lambda_p \setminus B_{i,j}\}\) for any \((i, j) \in \Lambda_p\). By Lemma 4.10,

\[
|P(V'_n \leq a_n) - e^{-\lambda_n}| \leq b_{1,n} + b_{2,n}
\]

where

\[
\lambda_n = |\Lambda_p| \cdot P(A_{1,\tau+1}), \quad b_{1,n} \leq \sum_{d \in \Lambda_p} \sum_{d' \in B_{a}} P(A_{1,\tau+1})^2 = 8\tau p^3 P(A_{1,\tau+1})^2
\]

and

\[
b_{2,n} \leq \sum_{d \in \Lambda_p} \sum_{d \neq d' \in B_{a}} P(Z_d > t, Z_{d'} > t)
\]

(4.44)
from the fact that \( \{Z_{ij}; \ (i,j) \in \Lambda_p\} \) are identically distributed.

We first calculate \( \lambda_n \). By definition of \( \Lambda_p \),

\[
\frac{p^2}{2} > |\Lambda_p| \geq \left| \{(i,j); \ 1 \leq i < j \leq p, \ j - i \geq \tau \} \right| - 2p \cdot |\Gamma_{p,\delta}|
\]

\[
= \sum_{i=1}^{p-\tau} (p - \tau - i + 1) - 2p \cdot |\Gamma_{p,\delta}|. \tag{4.45}
\]

Now the sum above is equal to \( \sum_{j=1}^{p-\tau} j = (p-\tau)(p-\tau+1)/2 \sim p^2/2 \) since \( \tau = o(p_t) \) for any \( t > 0 \). By assumption \( |\Gamma_{p,\delta}| = o(p) \) we conclude that

\[
|\Lambda_p| \sim \frac{p^2}{2} \tag{4.46}
\]

as \( n \to \infty \). It then follows from (4.41) that

\[
\lambda_n \sim \frac{p^2}{2} \cdot e^{-y/2} \cdot \frac{1}{p^2} \sim \frac{e^{-y/2}}{2}. \tag{4.47}
\]

as \( n \to \infty \).

Recall (4.43) and (4.47), to complete the proof, we have to verify that \( b_{1,n} \to 0 \) and \( b_{2,n} \to 0 \) as \( n \to \infty \). Clearly, by the first expression in (4.44), we get from (4.47) and then (4.46) that

\[
b_{1,n} \leq 8\tau p^3 P(A_{1\tau+1})^2 = \frac{8\tau p^3 \lambda_n}{|\Lambda_p|^2} = O\left(\frac{\tau}{p}\right) \to 0 \tag{4.48}
\]

as \( n \to \infty \) by the assumption on \( \tau \).

**Step 3.** Now we consider \( b_{2,n} \). Write \( d = (d_1, d_2) \in \Lambda_p \) and \( d' = (d_3, d_4) \in \Lambda_p \) with \( d_1 < d_2 \) and \( d_3 < d_4 \). It is easy to see that

\[
b_{2,n} \leq 2 \sum P(Z_d > a_n, Z_{d'} > a_n),
\]

where the sum runs over every pair \((d, d')\) satisfying

\[
d, d' \in \Lambda_p, \ d \neq d', \ d_1 \leq d_3 \text{ and } |d_i - d_j| < \tau
\]

for some \( i \in \{1, 2\} \) and some \( j \in \{3, 4\} \). \tag{4.49}
Geometrically, there are three cases for the locations of $d = (d_1, d_2)$ and $d' = (d_3, d_4)$:

\[ (1) \, d_2 \leq d_3; \quad (2) \, d_1 \leq d_3 < d_4 \leq d_2; \quad (3) \, d_1 \leq d_3 \leq d_2 \leq d_4. \quad (4.50) \]

Let $\Omega_j$ be the subset of index $(d, d')$ with restrictions (4.49) and $(j)$ for $j = 1, 2, 3$. Then

\[ b_{2,n} \leq 2 \sum_{i=1}^{3} \sum_{(d,d') \in \Omega_i} P(Z_d > a_n, Z_{d'} > a_n). \quad (4.51) \]

We next analyze each of the three sums separately. Recall all diagonal entries of $\Sigma$ in $N_p(0, \Sigma)$ are equal to 1. Let random vector

\[ (w_1, w_2, \ldots, w_p) \sim N_p(0, \Sigma), \quad (4.52) \]

then every $w_i$ has the distribution of $N(0, 1)$.

**Case (1).** Evidently, (4.49) and (1) of (4.50) imply that $0 \leq d_3 - d_2 < \tau$. Hence, $|\Omega_1| \leq \tau p^3$. Further, for $(d, d') \in \Omega_1$, the covariance matrix of $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ is equal to

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & \gamma & 0 \\
0 & \gamma & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]

for some $\gamma \in [-1, 1]$. Thus, the covariance matrix of $(w_{d_2}, w_{d_1}, w_{d_3}, w_{d_4})$ is equal to

\[
\begin{pmatrix}
1 & 0 & \gamma & 0 \\
0 & 1 & 0 & 0 \\
\gamma & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]
Recall $Z_d = Z_{d_1,d_2} = Z_{d_2,d_1} = 2[(\sum_{k=1}^{n/2} x_{kd_1} x_{kd_2})^2 + (\sum_{k=n/2+1}^{n} x_{kd_1} x_{kd_2})^2]$ defined at the beginning of Step 2. By Lemma 4.5, for some $\epsilon > 0$ small enough,

$$\sum_{(d,d') \in \Omega_1} P(Z_d > a_n, Z_{d'} > a_n) = \sum_{(d,d') \in \Omega_1} P(Z_{d_2,d_1} > a_n, Z_{d_3,d_4} > a_n)$$

$$\leq \tau p^3 \cdot O\left(\frac{1}{p^{1-\epsilon}}\right) = O\left(\frac{\tau}{p^{1-\epsilon}}\right) \to 0 \quad (4.53)$$

as $n \to \infty$ since $\tau = o(p^t)$ for any $t > 0$.

**Case (2).** For any $(d, d') \in \Omega_2$, there are three possibilities:

(I): $|d_1 - d_3| < \tau$ and $|d_2 - d_4| < \tau$; (II): $|d_1 - d_3| < \tau$ and $|d_2 - d_4| \geq \tau$; (III): $|d_1 - d_3| \geq \tau$ and $|d_2 - d_4| < \tau$. The case that $|d_1 - d_3| \geq \tau$ and $|d_2 - d_4| \geq \tau$ is excluded by (4.49).

Let $\Omega_{2,I}$ be the subset of $(d, d') \in \Omega_2$ satisfying (I), and $\Omega_{2,II}$ and $\Omega_{2,III}$ be defined similarly. It is easy to check that $|\Omega_{2,I}| \leq \tau^2 p^2$. The covariance matrix of $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ is equal to

$$\begin{pmatrix}
1 & 0 & \gamma_1 & 0 \\
0 & 1 & 0 & \gamma_2 \\
\gamma_1 & 0 & 1 & 0 \\
0 & \gamma_2 & 0 & 1
\end{pmatrix}$$

for some $\gamma_1, \gamma_2 \in [-1, 1]$. By Lemma 4.7,

$$\sum_{(d,d') \in \Omega_{2,I}} P(Z_d > a_n, Z_{d'} > a_n) = O\left(\frac{\tau^2}{p^{1-\epsilon}}\right) \to 0 \quad (4.54)$$

as $n \to \infty$.

Observe $|\Omega_{2,II}| \leq \tau p^3$. The covariance matrix of $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ is equal to

$$\begin{pmatrix}
1 & 0 & \gamma & 0 \\
0 & 1 & 0 & 0 \\
\gamma & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}, \quad |\gamma| \leq 1.$$
By Lemma 4.5, take $\epsilon > 0$ small enough to get
\[
\sum_{(d,d') \in \Omega_{2,III}} P(Z_d > a_n, Z_{d'} > a_n) = O\left(\frac{\tau}{p^{1-\epsilon}}\right) \to 0
\] as $n \to \infty$.

The third case is similar to the second one. In fact, $|\Omega_{2,III}| \leq \tau p^3$. The covariance matrix of $(w_{d_1}, w_{d_2}, w_{d_3}, w_{d_4})$ is equal to
\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & \gamma \\
0 & 0 & 1 & 0 \\
0 & \gamma & 0 & 1
\end{pmatrix}, \quad |\gamma| \leq 1.
\]

Thus, the covariance matrix of $(w_{d_2}, w_{d_1}, w_{d_4}, w_{d_3})$ is equal to $\Sigma_4$ in Lemma 4.5. Then, by the same argument as that in the equality in (4.53) we get
\[
\sum_{(d,d') \in \Omega_{2,III}} P(Z_d > a_n, Z_{d'} > a_n) = O\left(\frac{\tau}{p^{1-\epsilon}}\right) \to 0
\] as $n \to \infty$ by taking $\epsilon > 0$ small enough. Combining (4.54), (4.55) and (4.56), we conclude
\[
\sum_{(d,d') \in \Omega_2} P(Z_d > a_n, Z_{d'} > a_n) \to 0
\] as $n \to \infty$. This and (4.53) together with (4.51) say that, to finish the proof of this proposition, it suffices to verify
\[
\sum_{(d,d') \in \Omega_3} P(Z_d > a_n, Z_{d'} > a_n) \to 0
\] as $n \to \infty$. The next lemma confirms this. The proof is completed.  

**Lemma 4.8** Let the notation be as in the proof of Proposition 4.4. Then, for $\delta \in (0, 1)$, there exists $\epsilon_0 = \epsilon(\delta) > 0$ such that
\[
\sum_{(d,d') \in \Omega_3} P\left(Z_{12} > a_n, Z_{34} > a_n\right) = O\left(\frac{\tau}{p^{\epsilon_0}}\right)
\] as $n \to \infty$. 

\[\Box\]
The proof will appear in the section 4.4.

**Proof of Theorem 4.4.** By the first paragraph in the proof of Theorem 4.1, w.l.o.g., we prove the theorem by assuming that the \( n \) rows of \( X_n = (x_{ij})_{1 \leq i \leq n, 1 \leq j \leq p} \) are i.i.d. random vectors with distribution \( N_p(0, \Sigma) \) where all of the diagonal entries of \( \Sigma \) are equal to 1. Consequently, by the assumption on \( \Sigma \), for any subset \( E = \{i_1, i_2, \ldots, i_m\} \) of \( \{1, 2, \ldots, p\} \) with \( |i_s - i_t| \geq \tau \) for all \( 1 \leq s < t \leq m \), we know that \( \{x_{ki}; 1 \leq k \leq n, i \in E\} \) are \( mn \) i.i.d. \( N(0,1) \)-distributed random variables.

Noticing in the indice of \( \max_{1 \leq i < j \leq p, |i-j| \geq \tau} r_{ij}^2 = L_{n, \tau} \) and using the same idea as in Lemma 4.1, it is easy to see that

\[
\Delta_n := |n^2L_{n, \tau} - V_{n, \tau}|
\leq (B_1^2 + 2B_1)V_{n, \tau} + 4(B_1B_2 + B_2)\sqrt{V_{n, \tau}} + 4B_2^2.
\]

Reviewing the proof of Lemma 4.2, the argument is only based on the distribution of each column of \( \{x_{ij}\}_{n \times p} \); the joint distribution of any two different columns are irrelevant. In current situation, the entries in each column are i.i.d. standard normals. Thus, take \( \alpha = 2 \) in the lemma to have

\[
\{\sqrt{\frac{n}{\log p}}B_1\} \text{ and } \{\frac{B_2}{\log p}\} \text{ are tight,}
\]

when \( n \to \infty, p \to \infty \) with \( \log p = o(n^{1/3}) \), where \( B_1 \) and \( B_2 \) are as in Lemma 4.2. Let \( V_n = V_{n, \tau} \) be as in (4.38). It is seen from Proposition 4.4 that

\[
\frac{V_{n, \tau}}{n \log p} \to 4
\]

in probability as \( n \to \infty, p \to \infty \) and \( \log p = o(n^{1/3}) \). Now, using (4.59), (4.60) and (4.59), replacing \( W_n \) with \( V_{n, \tau} \) and \( L_n \) with \( L_{n, \tau} \) in the proof of Theorem 4.3, and repeating the whole proof again, we obtain

\[
\frac{n^2L_{n, \tau}^2 - V_{n, \tau}}{n} \to 0
\]
in probability as $n \to \infty$. This joint with Proposition 4.4 and the Slusky lemma yields the desired limiting result for $L_{n,\tau}$. ■

**Proof of theorem 4.5.** Set $a_n = 4n \log p + ny$, using the same idea in the proof of (4.42) in Proposition 4.4, we know that

$$|P(V_n \geq a_n) - P(V'_n \geq a_n)| \leq C\sqrt{\frac{(\log p)^3}{n}}$$

when $|\Gamma_{p,\delta}| \leq Cp\sqrt{\frac{(\log p)^3}{n}}$. Now we need to prove

$$|P(V'_n \geq a_n) - \exp\left(-\frac{1}{2}e^{-\frac{y}{2}}\right)| \leq C\sqrt{\frac{(\log p)^3}{n}}. \quad (4.61)$$

By lemma 4.10, we only need to get the values of $|\exp(-\lambda_n) - \exp(-\frac{1}{2}e^{-\frac{y}{2}})|$, $b_{1,n}$ and $b_{2,n}$.

By the equation (4.45), $|\Lambda_p - \frac{y^2}{2}| \leq C\sqrt{\frac{(\log p)^3}{n}}$ as $n \to \infty$, then follows from lemma 4.3 that $|\lambda_n - \frac{e^{-\frac{y}{2}}}{2}| \leq C\sqrt{\frac{(\log p)^3}{n}}$ as $n \to \infty$. So

$$|\exp(-\lambda_n) - \exp(-\frac{1}{2}e^{-\frac{y}{2}})| \leq C\sqrt{\frac{(\log p)^3}{n}} \quad (4.62)$$

as $n \to \infty$. Recall (4.48), we already have that $b_{1,n} \leq 8\tau p^3P(A_{1,\tau+1})^2 = \frac{8\tau p^3\lambda_n^2}{|\Lambda_p|^2} = O\left(\frac{z}{p}\right)$ as $n \to \infty$. Now combining (4.54), (4.55) and (4.56), and the lemma 4.8 , exit $\epsilon_0$

$$b_{2,n} \leq O\left(\frac{\tau}{p_0}\right).$$

Since $\frac{\log n}{\log p} \to 0$ and $\tau = o(p^t)$ for any $t > 0$, we know (4.61) is right. Under the condition $\frac{\log n}{\log p} \to 0$, we follow the exact steps in the proof of Theorem 4.3, and conclude

$$\left|P(nL_{n,\tau}^2 - 4n \log p \leq y) - \exp(-\frac{1}{2}e^{-\frac{y}{2}})\right| \leq C\sqrt{\frac{(\log p)^3}{n}}.$$  ■
4.4 Auxiliary Lemmas

In this section, we will prove Lemma 4.1-4.8 used in section 3. In here, if there is no confusion, we will use \( p \) instead of \( p_n \). And define \( \| \theta \| = (\sum_{i=1}^{n} \theta_i^2)^{\frac{1}{2}} \) for any \( n \)-dimensional vector \( \theta \).

**Proof of Lemma 4.1** By the definitions, it is easy to see that
\[
|n^2 r_{i,j}^2 - 2(A_{n,i,j}^2 + B_{n,i,j}^2)| \leq 2|\tilde{A}_{n,i,j}^2 h_i^2 h_j^2 | - A_{n,i,j}^2 | + 2|\tilde{B}_{n,i,j}^2 h_i^2 h_j^2 | - B_{n,i,j}^2 |, \tag{4.63}
\]
\[
|n^2 L_n^2 - W_n^2| \leq \max_{1 \leq i < j \leq p} |n^2 r_{i,j}^2 - 2(A_{n,i,j}^2 + B_{n,i,j}^2)|. \tag{4.64}
\]

First consider that
\[
\left| \frac{\tilde{A}_{n,i,j}^2}{h_i^2 h_j^2} - A_{n,i,j}^2 \right| = \left| \frac{\tilde{A}_{n,i,j}^2}{h_i^2 h_j^2} - A_{n,i,j}^2 \right| h_i^2 h_j^2 + A_{n,i,j}^2 |. \tag{4.65}
\]

It is easy to see that
\[
\left| \frac{\tilde{A}_{n,i,j}^2}{h_i^2 h_j^2} - A_{n,i,j}^2 \right| \leq |A_{n,i,j}^2| \left| \frac{1}{h_i^2 h_j^2} - 1 \right| + \left( \frac{n}{2} + 1 \right) \left( |\tilde{x}_i^{(1)}\tilde{x}_j| + |\tilde{x}_j^{(1)}\tilde{x}_i| + |\tilde{x}_i \tilde{x}_j| \right)
\]
\[
\leq |A_{n,i,j}^2| (2b_{n,1}^2 + 2b_{n,2}^2 + 2 + n)b_{n,2}^2 \leq |A_{n,i,j}^2| B_1 + B_2
\]

So
\[
\left| \frac{\tilde{A}_{n,i,j}^2}{h_i^2 h_j^2} - A_{n,i,j}^2 \right| \leq A_{n,i,j}^2 (B_1^2 + 2B_1) + 2(B_1 B_2 + B_2)|A_{n,i,j}^2| + B_2^2. \tag{4.65}
\]

Due to the \( i.i.d. \) assumption on \( x_{i,j} \), we get that
\[
\left| \frac{\tilde{B}_{n,i,j}^2}{h_i^2 h_j^2} - B_{n,i,j}^2 \right| \leq B_{n,i,j}^2 (B_1^2 + 2B_1) + 2(B_1 B_2 + B_2)|B_{n,i,j}^2| + B_2^2. \tag{4.66}
\]

By the definition of \( W_n^2 \), combine with (4.63), (4.64), (4.65) and (4.66), we can prove this lemma. □
Proof of Lemma 4.2. Recall that a sequence of random variables \( \{X_n; n \geq 1\} \) are said to be tight if, for any \( \epsilon > 0 \), there is a constant \( K > 0 \) such that 
\[
\sup_{n \geq 1} P(|X_n| \geq K) < \epsilon.
\]
Obviously, \( \{X_n; n \geq 1\} \) are tight if for some \( K > 0 \), 
\[
\lim_{n \to \infty} P(|X_n| \geq K) \to 0.
\]
From now on, we choose \( K > 3 \).

(i) First, since \( x_{ij} \)'s are i.i.d. bounded random variables with mean zero and variance one, by (i) of Lemma 4.12, when \( n \) is large enough,
\[
P\left( \frac{\sqrt{n}}{\log p} b_{n,3} \geq K \right) = P\left( \max_{1 \leq i \leq p} \left| \frac{1}{\sqrt{n \log p}} \sum_{k=1}^{n} x_{ki} \right| \geq K \right) 
\leq p \cdot P\left( \left| \frac{1}{\sqrt{n \log p}} \sum_{k=1}^{n} x_{k1} \right| \geq K \right) 
\leq p \cdot e^{-\left( \frac{K^2}{3} \right) \log p} = \frac{1}{p^{K^2/3-1}}. \tag{4.67}
\]
This says that \( \{\sqrt{n}/\log p b_{n,3}\} \) are tight.

Second, without loss of generality, we assume \( n \) is even. When \( n \) is large enough,
\[
P\left( \frac{\sqrt{n}}{\log p} b_{n,4} \geq 2K \right) = P\left( \max_{1 \leq i \leq p} \left| \sqrt{n \log p} \max(\bar{x}_{i}^{(1)}, \bar{x}_{i}^{(2)}) \right| \geq 2K \right) 
\leq 2p \cdot P\left( \left| \frac{1}{\sqrt{n \log p}} \sum_{k=1}^{n/2} x_{k1} \right| \geq K \right) 
\leq 2p \cdot e^{-\left( \frac{K^2}{3} \right) \log p} = \frac{2}{p^{K^2/3-1}}. \tag{4.68}
\]

Third, noticing that \( |t - 1| \leq |t^2 - 1| \) for any \( t > 0 \) and \( n h_i^2 = x_i^T x_i - n|\bar{x}_i|^2 \), we get that
\[
b_{n,1} \leq \max_{1 \leq i \leq p} |h_i^2 - 1| \leq \max_{1 \leq i \leq p} \left| \frac{1}{n} \sum_{k=1}^{n} (x_{ki}^2 - 1) \right| + \max_{1 \leq i \leq p} \left| \frac{1}{n} \sum_{k=1}^{n} x_{ki} \right|^2 
= Z_n + b_{n,4}^2,
\]
where \( Z_n = \max_{1 \leq i \leq p} \left| \frac{1}{n} \sum_{k=1}^{n} (x_{ki}^2 - 1) \right| \). Therefore, as \( n \to \infty \),
\[
P\left( \sqrt{\frac{n}{\log p}} b_{n,1} \geq 2K \right) \leq P\left( \sqrt{\frac{n}{\log p}} Z_n \geq K \right) + P\left( \sqrt{\frac{\log p}{n}} \left( \sqrt{\frac{n}{\log p}} b_{n,4} \right)^2 \geq K \right) 
\leq 2 \cdot \frac{1}{p^{K^2/3-1}} \to 0. \tag{4.69}
\]
Hence, we conclude that \( \{\sqrt{n/\log p}b_{n,1}\} \) are tight.

Fourth, when \( n \) is large enough, for any \( \epsilon > 0 \) and \( K > 0 \),

\[
P(|b_{n,2}^2 - 1| \geq \epsilon) \leq P(\max_{1 \leq i \leq p} |h_i^2 - 1| \geq \epsilon) \leq \frac{1}{p^{n\epsilon^2/(3\log p)-1}}. \tag{4.70}
\]

We know that \( n/\log p \to \infty \) and \( b_{n,2} \geq 0 \), then \( b_{n,2} \) converge to 1 in probability as \( n \to \infty \).

Finally, for any \( 1 > \epsilon > 0 \), when \( n \) is large enough, combining all the above inequalities, we get

\[
P(\sqrt{n/\log p}B_1 \geq 4K) \leq P(\sqrt{n/\log p}b_{n,1}^2 + 2b_{n,1}b_{n,2}^2 \geq 4K, |b_{n,2}^2 - 1| \leq \epsilon) + P(|b_{n,2}^2 - 1| \geq \epsilon) \leq \frac{2}{p^{K(1-\epsilon)^2/3-1}} \tag{4.71}
\]

and

\[
P(\frac{B_2}{\log p} \geq 8K) \leq P(\frac{n}{\log p}(b_{n,3}^2 + b_{n,4}^2) \geq 4K(1-\epsilon)) + P(|b_{n,2} - 1| \geq \epsilon) \leq \frac{3}{p^{K(1-\epsilon)^2/3-1}}. \tag{4.72}
\]

Then we know that \( \{\sqrt{n/\log p}B_1\} \) and \( \{B_2/\log p\} \) are tight.

(ii) Using \( a_n := \sqrt{\log p} = o(n^{\beta/2}) \) and (ii) of Lemma 4.12, we have

\[
P\left(\left|\frac{1}{\sqrt{n\log p}} \sum_{k=1}^n x_{k1}\right| \geq K\right) \leq \frac{1}{p^{K^2/3}} \quad \text{and}
\]

\[
P\left(\left|\frac{1}{\sqrt{n\log p}} \sum_{k=1}^n (x_{k1} - 1)\right| \geq K\right) \leq \frac{1}{p^{K^2/3}}
\]

as \( n \) is sufficiently large, where the first inequality holds provided \( E \exp(t_0|x_{11}|^{\alpha/2}) < \infty \); the second holds since \( E \exp(t_0|x_{11}^2 - 1|^{\alpha/2}) < \infty \) for some \( t_0 > 0 \), which is equivalent to \( E e^{t'0|x_{11}|^\alpha} < \infty \) for some \( t'_0 > 0 \). Then by the same method as in (i), we prove the same conclusions (4.67)-(4.72) under the assumption \( E \exp(t_0|x_{ij}|^\alpha) \leq \infty \).
Proof of Lemma 4.3. First we assume \( n \) is an even number. Take \( \gamma = \beta / \alpha_1 = (1 - \beta) / 2 \in [1/3, 1/2) \), set 
\[
\eta_i = \xi_i I(|\xi_i| \leq n^\gamma), \quad \mu_n = E\eta_1 \quad \text{and} \quad \sigma_n^2 = Var(\eta_1), \quad 1 \leq i \leq n. \tag{4.73}
\]
Since the desired result is a conclusion about \( n \to \infty \), without loss of generality, assume \( \sigma_n > 0 \) for all \( n \geq 1 \). We first claim that there exists a constant \( C > 0 \) such that
\[
\max \left\{ |\mu_n|, |\sigma_n - 1|, P(|\xi_1| > n^\gamma) \right\} \leq Ce^{-n^\beta/C} \tag{4.74}
\]
for all \( n \geq 1 \). In fact, since \( E\xi_1 = 0 \) and \( \alpha_1 \gamma = \beta \),
\[
|\mu_n| = |E\xi_1 I(|\xi_1| > n^\gamma)| \leq E|\xi_1|I(|\xi_1| > n^\gamma) \leq E\left(|\xi_1|e^{\alpha_1|\xi_1|/2}\right) \cdot e^{-\alpha_1n^\beta/2} \tag{4.75}
\]
for all \( n \geq 1 \). Note that \( |\sigma_n - 1| \leq |\sigma_n^2 - 1| = \mu_n^2 + E\xi_1^2 I(|\xi_1| > n^\gamma) \), by the same argument as in (4.75), we know both \( |\sigma_n - 1| \) and \( P(|\xi_1| > n^\gamma) \) are bounded by \( Ce^{-n^\beta/C} \) for some \( C > 0 \). Then (4.74) follows.

Step 1. We prove that, for some constant \( C > 0 \),
\[
|P\left( \frac{2(S_n^2 + T_n^2)}{n \log p_n} \geq y_n^2 \right) - P\left( \frac{2(\sum_{i=1}^{\frac{n}{2}} \eta_i)^2 + (\sum_{i=\frac{n}{2}+1}^{n} \eta_i)^2}{n \log p_n} \geq y_n^2 \right) | \leq Cne^{-n^\beta/C} \tag{4.76}
\]
for all \( n \geq 1 \). Observe
\[
\xi_i \equiv \eta_i \quad \text{for} \quad 1 \leq i \leq n \quad \text{if} \quad \max_{1 \leq i \leq n} |\xi_i| \leq n^\gamma, \tag{4.77}
\]
then by (4.74)
\[
P\left( \frac{2(S_n^2 + T_n^2)}{n \log p_n} \geq y_n^2 \right) \leq P\left( \frac{2(S_n^2 + T_n^2)}{n \log p_n} \geq y_n, \max_{1 \leq i \leq n} |\xi_i| \leq n^\gamma \right) + P\left( \bigcup_{i=1}^{n}\{|\xi_i| > n^\gamma\} \right) \leq P\left( \frac{2(\sum_{i=1}^{\frac{n}{2}} \eta_i)^2 + (\sum_{i=\frac{n}{2}+1}^{n} \eta_i)^2}{n \log p_n} \geq y_n^2 \right) + Cne^{-n^\beta/C} \tag{4.78}
\]
for all $n \geq 1$. Use inequality that $P(AB) \geq P(A) - P(B^c)$ for any events $A$ and $B$ to have

\[
P\left(\frac{2(S_n^2 + T_n^2)}{n \log p_n} \geq y_n^2\right) \geq P\left(\frac{2(S_n^2 + T_n^2)}{n \log p_n} \geq y_n^2, \max_{1 \leq i \leq n} |\xi_i| \leq n^\gamma\right)
\]

\[
= P\left(\frac{2([\sum_{i=1}^{n} \eta_i]^2 + ([\sum_{i=\frac{n}{2}+1}^{n} \eta_i])^2]}{n \log p_n} \geq y_n^2, \max_{1 \leq i \leq n} |\xi_i| \leq n^\gamma\right)
\]

\[
\geq P\left(\frac{2([\sum_{i=1}^{\frac{n}{2}} \eta_i]^2 + ([\sum_{i=\frac{n}{2}+1}^{n} \eta_i])^2]}{n \log p_n} \geq y_n^2\right) - Cn e^{-n^\beta/C}.
\]

This and (4.78) concludes (4.76).

**Step 2.** Set

\[
\eta_i' = \frac{\eta_i - \mu_n}{\sigma_n}
\]

for $1 \leq i \leq n$. First we will prove

\[
P\left(\frac{\sum_{i=1}^{\frac{n}{2}} \eta_i'}{\sqrt{n \log p_n}} \geq y_n'\right) \sim 1 - \Phi(z_n)
\]

as $n \to \infty$, where

\[
z_n = y_n' \sqrt{\log p_n} \quad \text{and} \quad y_n' = \frac{1}{\sigma_n} \left( y_n \pm \sqrt{\frac{n}{\log p_n}} \mu_n \right).
\]

By (4.74) and both $\sigma_n$ and $y_n$ have finite limits,

\[
|y_n' - y_n| \leq \frac{|1 - \sigma_n|}{\sigma_n} y_n + C \cdot \sqrt{\frac{n}{\log p_n}} |\mu_n| \leq C n e^{-n^\beta/C}
\]

for all $n \geq 1$. In particular, since $\log p_n = o(n^\beta)$,

\[
z_n = y_n' \sqrt{\log p_n} = o(n^{\beta/2})
\]

as $n \to \infty$. Reviewing (4.73), for some constant $K > 0$, we have $|\eta_i'| \leq Kn^\gamma$ for $1 \leq i \leq n$. Take $c_n = Kn^{\gamma-1/2} = Kn^{-\beta/2}$, recalling $z_n$ in (4.83), it is easy to check that as $n \to \infty$,

\[
s_n := \left(\sum_{i=1}^{\frac{n}{2}} E|\eta_i'|^2\right)^{1/2} = \sqrt{\frac{n}{2}}, \quad g_n := \sum_{i=1}^{\frac{n}{2}} E|\eta_i'|^3 \sim nC, \quad |\eta_i'| \leq c_n s_n \quad \text{and} \quad 0 < c_n \leq 1
\]
as $n$ is sufficiently large. Recall $\gamma = (1 - \beta)/2$, it is easy to see from (4.83) that

$$0 < z_n < \frac{1}{18c_n}$$

for $n$ large enough. By (4.81) and (4.82), $z_n s_n = y_n' \sqrt{n \log p_n/2}$ and $z_n \to \infty$ as $n \to \infty$. Use Lemma 4.11, it is easy to see

$$P\left( \frac{\sum_{i=1}^{n} \eta_i'}{\sqrt{n \log p_n/2}} \geq y_n' \right) = P\left( \sum_{i=1}^{n} \eta_i' \geq z_n s_n \right) = e^{r(z_n/s_n)}(1 - \Phi(z_n))(1 + \theta_{n,z_n}(1 + z_n)s_n^{-3} \varrho_n)

(4.84)$$

where $\vartheta_{n,z_n} = e^{r(z_n/s_n)} - 1 + e^{r(z_n/s_n)} \vartheta_{n,z_n}(1 + z_n)s_n^{-3} \varrho_n$ and $|\theta_{n,z_n}| \leq 36$. So (4.80) is correct. By the same idea,

$$P\left( \frac{\sum_{i=n+1}^{n/2} \eta_i'}{\sqrt{n \log p_n/2}} \geq y_n' \right) = P\left( \sum_{i=n+1}^{n} \eta_i' \geq z_n s_n \right) = (1 - \Phi(z_n))(1 + \vartheta_{n,z_n}).

(4.85)$$

Now, let $\gamma(x)$ be as in Lemma 4.11, since $\beta \leq 1/3$, combine with (4.83),

$$|\gamma\left(\frac{z_n}{s_n}\right)| \leq \frac{2z_n^3 \varrho_n}{s_n^3} = C \sqrt{\frac{(\log p_n)^3}{n}} = o\left(n^{3\beta - \frac{1}{2}}\right),$$

and

$$\frac{1 + z_n \varrho_n}{s_n^3} = O\left(n^{(\beta - 1)/2}\right),$$

as $n \to \infty$, and $|\theta_{n,x}| \leq 36$, then we conclude $|\vartheta_{n,z_n}| \leq C_n$, where $C_n \leq C \sqrt{(\log p_n)^3/n}$.

Now by the Lemma 4.13, we have

$$P\left(\frac{2(\sum_{i=1}^{n/2} \eta_i')^2 + 2(\sum_{i=n/2+1}^{n} \eta_i')^2}{n \log p_n} \geq (y_n')^2\right) = P(\chi^2(2) \geq z_n^2(1 + \vartheta_{n,z_n})),

(4.86)$$

where $|\vartheta_{n,z_n}| \leq 3C_n \to 0$ as $n \to \infty$.

**Step 3.** By the definition of (4.79) and let $b_n = \frac{n \log p_n}{2}$, we can have know that

$$\|b_n^{-\frac{1}{2}}(\sum_{i=1}^{n/2} \eta_i, \sum_{i=n/2+1}^{n} \eta_i')\| = \|b_n^{-\frac{1}{2}}\sigma_n(\sum_{i=1}^{n/2} \eta_i, \sum_{i=n/2+1}^{n} \eta_i') + b_n^{-\frac{1}{2}}(\mu_n, \mu_n)\|,

(4.87)$$
where \(b_n^{-\frac{1}{2}}(\mu_n, \mu_n)\) = \(\sqrt{\frac{n}{\log p_n}} \mu_n\). For any vector \(X\) and \(Y\), it is known that \(\|X\| - \|Y\| \leq \|X + Y\| \leq \|X\| + \|Y\|\). So we can get

\[
P(\|b_n^{-\frac{1}{2}}(\sum_{i=1}^{n/2} \eta_i, \sum_{i=n/2+1}^{n} \eta_i)\| \geq y_n)
\leq P(\|b_n^{-\frac{1}{2}}(\sum_{i=1}^{n/2} \eta_i', \sum_{i=n/2+1}^{n} \eta_i')\| \geq \frac{1}{\sigma_n} (y_n - \sqrt{\frac{n}{\log p_n}} u_n))
\]

and

\[
P(\|b_n^{-\frac{1}{2}}(\sum_{i=1}^{n/2} \eta_i, \sum_{i=n/2+1}^{n} \eta_i)\| \geq y_n)
\geq P(\|b_n^{-\frac{1}{2}}(\sum_{i=1}^{n/2} \eta_i', \sum_{i=n/2+1}^{n} \eta_i')\| \geq \frac{1}{\sigma_n} (y_n + \sqrt{\frac{n}{\log p_n}} u_n)).
\]

From the definition (4.81), we know \(z_n\) has two values \(z_{n1} = \frac{\sqrt{\log p_n}}{\sigma_n} (y_n - \sqrt{\frac{n}{\log p_n}} u_n)\) and \(z_{n2} = \frac{\sqrt{\log p_n}}{\sigma_n} (y_n + \sqrt{\frac{n}{\log p_n}} u_n)\). Combine the result (4.86),

\[
P(X^2(2) \geq z_{n2}^2)(1 + \vartheta'_{z_{n2},n}) \leq P(\|b_n^{-\frac{1}{2}}(\sum_{i=1}^{n/2} \eta_i, \sum_{i=n/2+1}^{n} \eta_i)\|^2 \geq y_n^2)
\leq P(X^2(2) \geq z_{n1}^2)(1 + \vartheta'_{z_{n1},n}).
\]

By the equation (4.82) and let \(t_n = \sqrt{\frac{n}{\log p_n}} \mu_n\), it is easy to see that

\[
P(X^2(2) \geq z_n^2) = \exp(-\frac{\gamma_n^2}{2}), |P(X^2(2) \geq z_n^2)\vartheta'_{z_n,n}| \leq C \sqrt{\frac{(\log p_n)^2}{n}} \frac{\gamma_n^2}{2} \cdot \frac{\gamma_n^2}{2}.
\]

and

\[
|\exp(-\frac{\gamma_n^2}{2} \log p_n) - \exp(-\frac{\gamma_n^2}{2} \log p_n)| \leq C \log p_n \exp^{-\gamma_n^2/C} \gamma_n^2 \cdot \frac{\gamma_n^2}{2} \cdot \frac{\gamma_n^2}{2}. \quad (4.88)
\]

When \(n\) is large enough, we have \(\log p_n \exp^{-\gamma_n^2/C} \gamma_n^2 \cdot \frac{\gamma_n^2}{2} \leq C \sqrt{\frac{(\log p_n)^2}{n} - \frac{\gamma_n^2}{2} \cdot \frac{\gamma_n^2}{2}}\). So
we can conclude that

\[
|P\left(\|b_n^{-\frac{1}{2}} \left( \sum_{i=1}^{n/2} \eta_i, \sum_{i=\frac{n}{2}+1}^{n} \eta_i \right) \|^2 \geq y_n^2 \right) - \exp\left(-\frac{y_n^2}{2} \log p_n\right)| \leq C \sqrt{\frac{(\log p_n)^3}{n} \frac{\delta_n^2}{n^2}}.
\]

(4.89)

Since \( t_n \leq C e^{-n^\beta/C} \), we have \( p_n^t_n = \exp(t_n \log p_n) \to 1 \) as \( n \to \infty \). Combining this with (4.76) and (4.89), we prove the Lemma for \( n \) is even.

Now, we consider \( n \) is an odd number. we know that \( S_n - \frac{1}{2} = S_n \) and \( T_n - \frac{1}{2} = T_n + \xi_n \). So \[
|P\left(\frac{2(S_n^2 + T_n^2)}{n \log p_n} \geq y_n^2 \right) - P\left(\frac{2(S_{n-1}^2 + T_{n-1}^2)}{n \log p_n} \geq y_n^2 - 16 \sqrt{\frac{\log p_n}{n}}\right)|
\leq P\left(\frac{T_n - \frac{1}{2} \xi_n}{n \log p_n} \geq 2 \sqrt{\frac{\log p_n}{n}}\right) + P\left(\frac{\xi_n^2}{n \log p_n} \geq 4 \sqrt{\frac{\log p_n}{n}}\right)
\leq C e^{-2(\log p_n)^3} + C e^{-C n^\beta},
\]

(4.90)

while the last inequality is by Lemma 4.12. Let \( d_n^2 = y_n^2 - 16 \sqrt{\frac{\log p_n}{n}} \),

\[
|P\left(\frac{2(S_{n-1}^2 + T_{n-1}^2)}{n \log p_n} \geq y_n^2 - 16 \sqrt{\frac{\log p_n}{n}}\right) - \exp\left(-\frac{y_n^2}{2} \log p_n\right)|
\leq C \sqrt{\frac{(\log p_n)^3}{n} \frac{\delta_n^2}{n^2}} + \left| \exp\left(\frac{y_n^2 - 16 \sqrt{\frac{\log p_n}{n}} \log p_n}{2}\right) - \exp\left(-\frac{y_n^2}{2} \log p_n\right)\right|
\leq C \sqrt{\frac{(\log p_n)^3}{n} \frac{\delta_n^2}{n^2}} + C \sqrt{\frac{(\log p_n)^3}{n} \frac{\delta_n^2}{n^2}},
\]

(4.91)

when \( n \) is large enough. Now we can combine the equations (4.90) and (4.91), and use the same proof as in (4.88), then the lemma is also right for \( n \) is odd. ■

**Proof of Lemma 4.4.** Since that \( A_{n,i,j}^2 = A_{n-1,i,j}^2 \), \( B_{n,i,j}^2 = B_{n-1,i,j}^2 + 2B_{n-1,i,j}x_{ni}x_{nj} \)
where \(E\), \(\Psi\), \(\eta\), \(n\)

So we can write \(\Psi\) is even from now.

while the last inequality is given by MDP. Without loss of generality, we assume \(n\) is even from now.

Then we can define the following random vectors

\[
\eta_k = \left( \frac{X_{k1}X_{k2}}{\sqrt{n \log p}}, \frac{X_{(k+\frac{1}{2})1X_{(k+\frac{1}{2})2}}}{\sqrt{n \log p}} \right), \quad \zeta_k = \left( \frac{X_{k1}X_{k3}}{\sqrt{n \log p}}, \frac{X_{(k+\frac{1}{2})1X_{(k+\frac{1}{2})2}}}{\sqrt{n \log p}} \right)
\]

So we can write \(\Psi_n = P\left( \| \sum_{k=1}^{n/2} \eta_k \| > t_n, \| \sum_{k=1}^{n/2} \zeta_k \| > t_n \right)\). Then

\[
\Psi_n \leq P\left( \| \sum_{k=1}^{n/2} (\eta_k + \zeta_k) \| > 2t_n \right) + P\left( \| \sum_{k=1}^{n/2} (\eta_k - \zeta_k) \| > 2t_n \right). \tag{4.93}
\]

Consider \(\eta_k + \zeta_k = \left( \frac{X_{k1}(X_{k2}+X_{k3})}{\sqrt{n \log p}}, \frac{X_{(k+\frac{1}{2})1(X_{(k+\frac{1}{2})2}+X_{(k+\frac{1}{2})3})}}{\sqrt{n \log p}} \right)\), define \(\xi_k = \frac{X_{k1}(X_{k2}+X_{k3})}{\sqrt{n \log p}}\), where \(E\xi_k = 0\), \(Var(\zeta_k) = 1\) and \(E\exp\{\alpha_1\zeta_k\} < \infty\). So apply Lemma 4.3,

\[
P\left( \| \sum_{k=1}^{n/2} (\eta_k + \zeta_k) \| > 2t_n \right) = P\left( \frac{2\left( \sum_{k=1}^{n/2} \xi_k \right)^2 + \left( \sum_{k=1}^{n/2} \xi_k \right)^2}{n \log p} > 2t_n^2 \right)
\sim \exp(-t_n^2 \log p_n)
\]

Similarly, the same result holds for \(P\left( \| \sum_{k=1}^{n/2} (\eta_k - \zeta_k) \| > 2t_n \right)\). Therefore, \(\Psi_n = O(P^{t^2-\epsilon})\). \(\blacksquare\)
Proof of Lemma 4.5, 4.6 and 4.7. Recall the definitions of \( Z_{12} \) and \( Z_{34} \):

\[
Z_{12} = 2\left(\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} u_{k1}u_{k2}\right)^2 + 2\left(\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} u_{k1}u_{k2}\right)^2;
\]

\[
Z_{34} = 2\left(\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} u_{k3}u_{k4}\right)^2 + 2\left(\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} u_{k3}u_{k4}\right)^2. \tag{4.94}
\]

Without loss of generality, we assume \( n \) is even. Define \( \eta_k = (u_{k1}u_{k2}, u_{(k+\frac{n}{2})1}u_{(k+\frac{n}{2})2}) \) and \( \zeta_k = (u_{k3}u_{k4}, u_{(k+\frac{n}{2})3}u_{(k+\frac{n}{2})4}) \), then \( Z_{12} = 2\|\sum_{k=1}^{n} \eta_k\|^2 \) and \( Z_{34} = 2\|\sum_{k=1}^{n} \zeta_k\|^2 \).

Then we have that

\[
P\left(Z_{12} > a_n, Z_{34} > a_n\right)
= P\left(2\|\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} \eta_k\|^2 > a_n, 2\|\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} \eta_k\|^2 > a_n\right)
\leq P\left(\|\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} (\eta_k + \zeta_k)\| > \sqrt{2a_n}\right) + P\left(\|\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} (\eta_k - \zeta_k)\| > \sqrt{2a_n}\right)
\leq P\left(\left(\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} (u_{k1}u_{k2} + u_{k3}u_{k4})\right)^2 + \left(\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} (u_{k1}u_{k2} + u_{k3}u_{k4})\right)^2 > 2a_n\right)
+ P\left(\left(\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} (u_{k1}u_{k2} - u_{k3}u_{k4})\right)^2 + \left(\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} (u_{k1}u_{k2} - u_{k3}u_{k4})\right)^2 > 2a_n\right).
\]

Under the conditions of Lemma 4.5 and 4.6, \( E(u_{k1}u_{k2} + u_{k3}u_{k4}) = 0 \) and \( Var(u_{k1}u_{k2} + u_{k3}u_{k4}) = 2 \). By the lemma 4.3,

\[
P\left(\left(\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} (u_{k1}u_{k2} + u_{k3}u_{k4})\right)^2 + \left(\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} (u_{k1}u_{k2} + u_{k3}u_{k4})\right)^2 > 2a_n\right) \sim O(p^{-4+\epsilon}),
\]

\[
P\left(\left(\sum_{k=1}^{\left\lfloor \frac{n}{2} \right\rfloor} (u_{k1}u_{k2} - u_{k3}u_{k4})\right)^2 + \left(\sum_{k=\left\lfloor \frac{n}{2} \right\rfloor + 1}^{n} (u_{k1}u_{k2} - u_{k3}u_{k4})\right)^2 > 2a_n\right) \sim O(p^{-4+\epsilon}).
\]

So \( P\left(Z_{12} > a_n, Z_{34} > a_n\right) \sim O(p^{-4+\epsilon}) \).
When under the condition of Lemma 4.7, \( \text{Var}(u_{k1}u_{k2} + u_{k3}u_{k4}) = 2 + 2\gamma_1\gamma_2. \)

By the lemma 4.3,

\[
P\left(\left(\sum_{k=1}^{n/2} (u_{k1}u_{k2} + u_{k3}u_{k4})\right)^2 + \left(\sum_{k=n/2+1}^{n} (u_{k1}u_{k2} + u_{k3}u_{k4})\right)^2 > 2a_n\right) \sim O(p^{-2(1+\epsilon)})
\]

\[
P\left(\left(\sum_{k=1}^{n/2} (u_{k1}u_{k2} - u_{k3}u_{k4})\right)^2 + \left(\sum_{k=n/2+1}^{n} (u_{k1}u_{k2} - u_{k3}u_{k4})\right)^2 > 2a_n\right) \sim O(p^{-2(1+\epsilon)}).
\]

So \( P\left(Z_{12} > a_n, Z_{34} > a_n\right) \sim O(p^{-2(1+\epsilon)}). \)

**Proof of Lemma 4.8.** Reviewing notation \( \Omega_3 \) defined below (4.50), the current case is that \( d_1 \leq d_3 \leq d_2 \leq d_4 \) with \( d = (d_1, d_2) \) and \( d' = (d_3, d_4) \). Of course, by definition, \( d_1 < d_2 \) and \( d_3 < d_4 \). To save notation, define the “neighborhood” of \( d_i \) as follows:

\[
N_i = \left\{ d \in \{1, \ldots, p\}; |d - d_i| < \tau \right\}
\]

for \( i = 1, 2, 3, 4 \). Given \( d_1 < d_2 \), there are two possibilities for \( d_4 \): (a) \( d_4 - d_2 > \tau \) and (b) \( 0 \leq d_4 - d_2 \leq \tau \). There are four possibilities for \( d_3 \): (A) \( d_3 \in N_2 \setminus N_1 \); (B) \( d_3 \in N_1 \setminus N_2 \); (C) \( d_3 \in N_1 \cap N_2 \); (D) \( d_3 \notin N_1 \cup N_2 \). There are eight combinations for the locations of \( (d_3, d_4) \) in total. However, by (4.49) the combination (a) & (D) is excluded. We can deal with all of the seven possibilities with the exact method as in the Lemma 6.13 in Cai and Jiang (2011), then the lemma is proved. \( \blacksquare \)

The following Lemmas which appear in the others’ paper, are the technical tools for us to prove the above lemmas and theorems.

The first result is the Bernstein dimension free inequality for independent random vector, we can find it from Prokhorov, A.V. (1968).

**Lemma 4.9** Let \( \xi_1, \xi_2, \ldots, \xi_n \) be i.i.d. vectors in \( R^2 \) such that \( E\xi_j = 0, \|\xi_j\| \leq L \). Let \( \lambda_1, \lambda_2 \) denote the eigenvalues of the covariance matrix of \( \xi_j \). When \( \lambda_1 = \lambda_2 \),

\[
P(\|\sum_{j=1}^{n} \xi_j\| \geq r\sqrt{n}) \leq \sqrt{\frac{\pi r^2}{2\lambda_1}} \exp \left\{ -\frac{r^2}{2\lambda_1} (1 + \frac{a}{3})^{-1} \right\}.
\]
where \( a = Lr/\lambda \sqrt{n} \), \( 0 \leq a \leq 1 \) and \( \frac{r^2}{2\lambda_1} \geq 3 \).

The following Poisson approximation result is essentially a special case of Theorem 1 from Arratia et al. (1989).

**Lemma 4.10** Let \( I \) be an index set and \( \{ B_\alpha, \alpha \in I \} \) be a set of subsets of \( I \), that is, \( B_\alpha \subset I \) for each \( \alpha \in I \). Let \( \{ \eta_\alpha, \alpha \in I \} \) be random variables. For a given \( t \in \mathbb{R} \), set \( \lambda = \sum_{\alpha \in I} P(\eta_\alpha > t) \). Then

\[
|P(\max_{\alpha \in I} \eta_\alpha \leq t) - e^{-\lambda}| \leq (1 + \lambda^{-1})(b_1 + b_2 + b_3)
\]

where

\[
b_1 = \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} P(\eta_\alpha > t)P(\eta_\beta > t),
\]

\[
b_2 = \sum_{\alpha \in I} \sum_{\alpha \neq \beta \in B_\alpha} P(\eta_\alpha > t, \eta_\beta > t),
\]

\[
b_3 = \sum_{\alpha \in I} E[P(\eta_\alpha > t|\sigma(\eta_\beta, \beta \notin B_\alpha)) - P(\eta_\alpha > t)],
\]

and \( \sigma(\eta_\beta, \beta \notin B_\alpha) \) is the \( \sigma \)-algebra generated by \( \{ \eta_\beta, \beta \notin B_\alpha \} \). In particular, if \( \eta_\alpha \) is independent of \( \{ \eta_\beta, \beta \notin B_\alpha \} \) for each \( \alpha \), then \( b_3 = 0 \).

The following conclusion is Example 1 from Sakhanenko (1991). See also Lemma 6.2 from Liu, Lin and Shao (2008).

**Lemma 4.11** Let \( \xi_i, 1 \leq i \leq n \), be independent random variables with \( E\xi_i = 0 \). Put

\[
s_n^2 = \sum_{i=1}^{n} E\xi_i^2, \quad g_n = \sum_{i=1}^{n} E|\xi_i|^3, \quad S_n = \sum_{i=1}^{n} \xi_i.
\]

Assume \( \max_{1 \leq i \leq n} |\xi_i| \leq c_n s_n \) for some \( 0 < c_n \leq 1 \). Then

\[
P(S_n \geq xs_n) = e^{\gamma(x/s_n)}(1 - \Phi(x))(1 + \theta_{n,x}(1 + x)s_n^{-3}g_n)
\]

for \( 0 < x \leq 1/(18c_n) \), where \( |\gamma(x)| \leq 2x^3g_n \) and \( |\theta_{n,x}| \leq 36 \).
The following are moderate deviation results from Chen (1990), see also Chen (1991), Dembo and Zeitouni (1998) and Ledoux (1992). They are a special type of large deviations.

**Lemma 4.12** Suppose $\xi_1, \xi_2, \ldots$ are i.i.d. r.v.’s with $E\xi_1 = 0$ and $E\xi_1^2 = 1$. Put $S_n = \sum_{i=1}^{n} \xi_i$.

(i) Let $0 < \alpha \leq 1$ and $\{a_n; n \geq 1\}$ satisfy that $a_n \to +\infty$ and $a_n = o(n^{\frac{\alpha}{2(2-\alpha)}})$. If $Ee^{t|\xi_1|\alpha} < \infty$ for some $t_0 > 0$, then
\[
\lim_{n \to \infty} \frac{1}{a_n^2} \log P\left(\frac{S_n}{\sqrt{na_n}} \geq u\right) = -\frac{u^2}{2}
\] (4.96)
for any $u > 0$.

(ii) Let $0 < \alpha < 1$ and $\{a_n; n \geq 1\}$ satisfy that $a_n \to +\infty$ and $a_n = O(n^{\frac{\alpha}{2(2-\alpha)}})$. If $Ee^{t|\xi_1|\alpha} < \infty$ for all $t > 0$, then (4.96) also holds.


**Lemma 4.13** Let $U_1, U_2, V_1$ and $V_2$ be independent random variables. Assume that there exist $0 \leq c_0 \leq 1$ and $x_0$ such that, for any $0 \leq x \leq x_0$,
\[
P(|U_1| \geq x) = P(|V_1| \geq x)(1 + \theta_{1,x}),
\]
\[
P(|U_2| \geq x) = P(|V_2| \geq x)(1 + \theta_{2,x}),
\]
where $|\theta_{1,x}| \leq c_0$ and $|\theta_{2,x}| \leq c_0$. Then
\[
P(U_1^2 + U_2^2 \geq x^2) = P(V_1^2 + V_2^2 \geq x^2)(1 + \theta_x)
\]
for $0 \leq x \leq x_0$, where $|\theta_x| \leq 3c_0$.

These two lemma are showed in Cai, T. Jiang, T (2011).

**Lemma 4.14** Suppose $\{x_{ij}; i \geq 1, j \geq 1\}$ be i.i.d. random variables with $|x_{11}| \leq C$ for a finite constant $C > 0$, $E|x_{11}| = 0$ and $E(x_{11}^2) = 1$. Assume $p = p(n) \to \infty$ and $\log p = o(n)$ as $n \to \infty$. Define $H_n = \max_{1 \leq i < j \leq p} |x_{ij}^T x_j| = \max_{1 \leq i < j \leq p} |\sum_{k=1}^{n} x_{ki} x_{kj}|$.

Then
\[
\frac{H_n}{\sqrt{n \log p}} \to 2
\]
in probability as $n \to \infty$. 

LEMMA 4.15 Suppose \( \{x_{ij}; i \geq 1, j \geq 1\} \) be i.i.d. random variables with \( E x_{11} = 0, E(x_{11}^2) = 1 \) and \( E e^{t|x_{11}|^\alpha} < \infty \). Assume \( p = p(n) \to \infty \) and \( \log p = o(n^\beta) \) as \( n \to \infty \), where \( \beta = \alpha/4 + \alpha \). Define \( H_n = \max_{1 \leq i < j \leq p} |x_i^T x_j| = \max_{1 \leq i < j \leq p} |\sum_{k=1}^n x_{ki} x_{kj}| \). Then

\[
\frac{H_n}{\sqrt{n \log p}} \to 2
\]

in probability as \( n \to \infty \).
Chapter 5

Discussion

In this thesis, we study the random matrix theory and its applications in three different types of random matrices: the truncated Haar unitary matrix, the beta-laguerre ensemble and the sample correlation matrix. Beyond this, we would like to bring out the following interesting open problems for future studies:

1. In chapter 2, we study the upper corner of a Haar unitary matrix. Similar results should also be true for Haar orthogonal matrices. In fact, Theorem 5 from Jiang (2009) shows that Theorem 2.2 in Section 1 is still valid if $U_n$ is replaced by a Haar orthogonal matrix and $m = o(\sqrt{n})$. Khoruzhenko, Sommers and Życzkowski (2010) and Forrester (2010) studied the density function of the eigenvalues of a truncated block of a Haar orthogonal matrix, which is shown to be not as clean as the case of a Haar unitary matrix. Therefore some additional work can be done in regard to Haar orthogonal matrices.

2. In chapter 2, we use the properties of the Ginibre ensembles to derive the circular law. This application brings us attention to this special type of random matrices, see Hough, et al. (2009). We like to consider the limiting spectral distributions and the large deviation principles of the product of $n$ complex non-Hermitian, independent random matrices, each of size
$N \times N$ with independent identically distributed Gaussian entries, even when $n \to \infty$ as $N \to \infty$. The joint density functions of the eigenvalues of these random matrices can be found in the paper of Akerman and Burda (2012). Furthermore, Krishnapar (2009) provided the joint density function for the eigenvalues of the $G_1G_2^{-1}$, where $G_1$ and $G_2$ are independent Ginibre matrices. We also like to provide the large deviation principles for this case too.

3. Random polynomial is a very interesting area that has a close relationship with random matrix theory. In some cases, the zeros of random polynomials are corresponding to the eigenvalues of random matrices, and the joint density function for these zeros exists in explicit form. Based on these results, we can treat the zeros of the random polynomial in the same way as the eigenvalues of a random matrix and derive the limiting empirical distributions of the zeros, the limiting distributions of the largest roots and so on. For example, Let $(Z_1, \cdots, Z_n)$ have density

$$C \prod_{1 \leq j<k \leq n} |z_j - z_k|^2 \cdot \prod_{j=1}^{n} \phi(z_j) \quad (5.1)$$

where $C$ is the normalizing constant and $\phi(x) \geq 0$ for all $x \geq 0$. We like to investigate the statistical properties of $(Z_1, \cdots, Z_n)$, such as the limiting distribution of the max

4. In Chapter 3, we do not derive the CLT for the spectral statistics of the sample covariance matrix when $p > n$. Such CLT can be used for deriving the limiting distribution of the test statistics for testing the independent structure of the covariance matrix in the case that the likelihood ratio test doesn’t work. Regarding the testing problems, although for testing sphericity, John (1972) proposed an alternative test statistic that is applicable when $p > n$ and he also derived a Chi-square approximation assuming fixed $p$. Ledoit and Wolf (2002) and Chen, et al. (2010) showed that John’s result remains
valid even for high-dimensional cases. For testing equality of multiple covariance matrices, Jiang and Yang (2013) proposed LRT for the case $p < n$. In the case that for $p > n$, the Wald statistic is used to perform the test. Schott (2007) further studied this statistic and gave its asymptotic null distribution for both fixed $p$ and large $p$ situations. Despite these enlightening work mentioned above, other hypothesis tests for $p > n$ are still an open area with many interesting problems to be solved.

5. In Chapter 4, we study the limiting distribution of the testing statistics, and use them to test the covariance structure. However we assume that exponential moment exist which is a strong condition. One problem we like to investigate is whether there is any statistics doesn’t depend on the structure of the population, such as sub-Gaussian tail. Furthermore, it is well known that the sample correlation coefficient $\rho = 0$ doesn’t mean independence. This is why the result is not suitable for testing the independence covariance structure without the normality assumption. Therefore, we can consider the Spearman rank correlation coefficient, which is defined in the following way: for a column $X_i = (X_{1i}, \cdots, X_{ni})^T$, we order the elements $X_{1i}, \cdots, X_{ni}$ of $X_i$ from least to greatest, and let $Q_{ki}$ denote the rank of $X_{ki}$ replacing the matrix $(X_{ki})$ by $(Q_{ki})$. The Spearman rank correlation coefficient between $X_i$ and $X_j$ is defined by

$$r_{ij} = \frac{12}{n(n^2 - 1)} \sum_{k=1}^{n} Q_{ki}Q_{kj} - \frac{3(n + 1)}{n - 1}. \quad (5.2)$$

The Spearman’s rank correlation matrix is given by

$$R_n = (r_{ij})_{1 \leq i,j \leq p}. \quad (5.3)$$

An import feature here is that the distribution of $r_{ij}$ does not depend on the distribution of $x_{ij}$. When $p$ and $n$ is comparable, and the $X_{ij}$ are i.i.d., Zhou (2007) already obtained the limiting distribution for the largest off-diagonal entry of $R_n$ with

$$l_n = \max_{1 \leq i,j \leq p} \| (r_{ij}) \|. \quad (5.4)$$
Deriving the limiting distribution for $l_n$ when $p$ is much larger than $n$, even when $p$ is less than $n$ in the exponential scale. Also, we like to consider the limiting distribution of $l'_n = \max_{1 \leq i, j \leq p, |i-j| \leq t} |r_{ij}|$, when $X_{11}, \cdots, X_{1p}$ have a banded covariance matrix with bandwidth equal to $t$. This statistic $l'_n$ can be used to test the banded covariance structure.
References


