Linköping Studies in Science and Technology
Dissertation No. 2313

Large Deviations of Condition Numbers and Extremal Eigenvalues of Random Matrices

Denise Uwamariya

LINKÖPING UNIVERSITY

Linköping Studies in Science and Technology.
Dissertations, No. 2313

Large deviations of condition numbers and extremal eigenvalues of random matrices

Denise Uwamariya

Linköping University
Department of Mathematics
Division of Mathematical Statistics
SE-581 83 Linköping, Sweden

Linköping 2023

Linköping Studies in Science and Technology. Dissertations, No. 2313

Large deviations of condition numbers and extremal eigenvalues of random matrices

Denise Uwamariya
Division of Mathematical Statistics
Department of Mathematics
Linköping University
SE–581 83 Linköping, Sweden
denise.uwamariya@liu.se
https://doi.org/10.3384/9789180751575

ISSN 0345-7524
ISBN 978-91-8075-156-8 (Print), 978-91-8075-157-5 (PDF)

Typeset by the author using LaTeX2e documentation system.

Printed by LiU-Tryck, Linköping, Sweden 2023

ABSTRACT

"Denise Uwamariya" (2023). Large deviations of condition numbers and extremal eigenvalues of random matrices. Doctoral dissertation 2313. ISBN 978-91-8075-156-8 (Print), 978-91-8075-157-5 (PDF). ISSN 0345-7524.

Random matrix theory found applications in many areas, for instance in statistics random matrices are used to analyse multivariate data and their eigenvalues are used in hypothesis testing. Spectral properties of random matrices have been studied extensively in the literature dealing with both the bulk case (involving all the eigenvalues) and the extremal case (addressing the maximal and minimal eigenvalues). In this thesis two types of sequences of random matrices are considered: the first type is the sequence of sample covariance matrices and the second type is the sequence of $\beta-$Laguerre (or Wishart) ensembles, for which large deviations of their extremal cases are studied. These two types of sequences of random matrices contain the classical Wishart matrices.

The thesis can be divided into two parts. The first part is on the study of large deviations of condition numbers defined as ratios of maximal and the minimal eigenvalues. This is done based on suitable analysis and estimates of the joint density function of all eigenvalues. The second part deals with large deviations of individual maximal and minimal eigenvalue, and the approach consists of suitable eigenvalue concentration inequalities and Laplace's method.

It is remarked that for those two types of sequences of random matrices considered in this thesis, two scenarios are investigated: either one of the dimension size and the sample size is much larger than the other one, or the two sizes are comparable.

Populärvetenskaplig sammanfattning

Slumpmatriser tillämpas inom många områden, till exempel, multivariat statistik och kärnfysik. Vi studerar slumpmatriser utifrån två aspekter: stokastiska konditionstal och extrema egenvärden.

Stokastiska konditionstal uppstod naturligt i numerisk linjär algebra på 1980-talet. Till exempel, för att uppskatta numerisk stabilitet och tillförlitlighet, behöver man ofta uppskatta sannolikheten att konditionstal är stora, och det här är vad vi fokuserar på — med hjälp av asymptotik studeras stora avvikelser av stokastiska konditionstal. Samtidigt spelar extrema egenvärden en viktig roll i statistisk referens. Till exempel kan det maximala egenvärdet användas för att testa strukture hos kovariansmatrisen för en slumpvektor.

Trots många års studier av stokastiska konditionstal och på grund av komplexitet med en täthets/sannolikhets funktion av extrema egenvärden och brist på lämpliga metoder, är stora avvikelser av stokastiska konditionstal och extrema egenvärden inte helt beskrivna i litteraturen.

Tidigare, när det var nödvändigt att uppskatta stora avvikelser sannolikheter för konditionstal och extrema egenvärden, baserades de vanligtvis på empiriska simuleringar. I denna avhandling föreslår vi nya metoder för att beskriva asymptotik för stora avvikelser av stokastiska konditionstal och extrema egenvärden.

Acknowledgments

I would like to take this great opportunity to thank every one who helped me in one way or the other during my PhD studies at Linköping University (LiU).

I am deeply grateful to my supervisor Xiangfeng Yang for all suggestions, advice, support, guidance and comments that lead to this thesis. Thank you for always leaving the door open for discussions.

Thank you my second supervisors Martin Singull and Dietrich von Rosen for corrections, comments and advise. I have been like to be part of the team with a good research atmosphere.

Thank you Swedish International Development Cooperation Agency (Sida) through Rwanda-Sweden bilateral programme for providing the financial support.

I would like to thank all lecturers at LIU, MAI department, administrative staffs, Lindha Nilsson your support is much appreciated. Thank you to my colleagues, current and former PhD students especially my office-mate Emelyne who I shared office all those years, Vincent for the help I received from him and Cigdem, Roghayeh, Jennifer, thank you for your great friendship.

Let me extend my thanks to LinStat'2022, LinStat'2020 and MMLM 2019 team organizers for great opportunity to attend the conferences and present my research work, I learned a lot as a young scientist and many thanks to LinStat'2020 for awarded prizes as the best PhD student presenter in the conference. Thank you also to the organizers of Junior female researchers in probability 2022 for giving me the opportunity to attend the workshop and financial support, it was great to meet and discuss with you strong women. Here, I acknowledge Zhenxia for sharing information about the workshop with me.

I would like to thank the University of Rwanda (UR), department of mathematics for the help I got to find balance between teaching and my studies. Finally, I want to thank my husband, Billy Second and our son, Derwin. We did it!

Thank you to all my family members and friends for the love and for believing in me.

Thank you all.

Linköping, April 20, 2023

"Denise Uwamariya"

Contents

Part I

Introduction, historical and theoretical background

1 Introduction

Random matrices play an important role in many different areas. As example, one can mention areas like statistics, physics, numerical analysis, neuroscience and many other areas. In the study of random matrices, as matrices whose entries are random variables chosen from some distribution, a major part is to study and understand the behavior of eigenvalues of square random matrices based on the distribution of the entries. The study can be summarized into two classes: i.e. the global (or the bulk) and the local (or the extremal) classes. The first class involves the study of all the eigenvalues, and the second class deals with the extremal eigenvalues (maximal and minimal eigenvalues).

Nowadays, random matrices usually have large sizes and in most cases it becomes difficult to apply classical results. This thesis as well as the appended papers provide asymptotic results of extremal eigenvalues of two types of sequences of random matrices: the sequence of sample covariance matrices and the sequence of $\beta-$Laguerre ensembles.

1.1 Historical background on random matrices

In the work around 1950 [59] Eugene Wigner studied square matrices of larger size, with applications in nuclear physics. Since then, the development of random matrices started to grow in many other areas. In multivariate statistics application, the development continues by the work of Constantine [7] in 1963, James [27, 28] in (1960, 1964), and Muirhead [35] in 1982. In nuclear physics one can cite the work by Wigner [57, 58, 60] in (1955, 1958), Gaudin [21] in 1956, Gurevich [24] in 1956, and Dyson [15, 16] in 1962. Afterwards, a lot of people worked on random matrices in areas beyond multivariate statistics and nuclear physics areas, for example numerical analysis [1, 11, 17], graph theory [6, 19, 18, 34], combinatorics [12, 9, 22, 26, 36, 44], engineering sciences [32, 45, 48, 51, 56, 62] and many others.

Random matrices were used for different purposes in the above-mentioned work, and one of the main research directions in the literature is the study of their eigenvalues as they play important roles. For instance, from derivations and technical results found in the books written by Muirhead 1982 [35] and Kollo and von Rosen 2005 [33], we can see that in statistics random matrices play a significant role in the analysis of multivariate data and several traditional tests use eigenvalues of the random matrices. In the Principal Component Analysis (PCA) first introduced in [39], as a statistical tool for data dimensional reduction, maximal eigenvalue of covariance matrix can be used to reduce the dimensional of the data and to check the structure of the covariance matrix given the data (see [42]). In Chapter 2, we present some basic concepts about sample covariance matrices and $\beta-$Laguerre ensembles. In Chapter 4 we summarize some known studies related to the spectral properties of these types of random matrices. Chapter 3 contains some theoretical background, while Chapter 5 includes the summary of results of the appended papers.

The focus of the thesis is the study of extremal eigenvalues of random matrices. In this thesis, we provide asymptotic results of individual maximal and minimal eigenvalues and condition numbers as the ratios of maximal and minimal eigenvalues of sample covariance matrices and $\beta-$Laguerre ensembles in terms of large deviations.

1.2 Aims

The overall aim of this thesis is to study asymptotics, in terms of large deviations, for the extremal eigenvalues and the condition numbers of sample covariance matrices and $\beta-$Laguerre ensembles when one of the dimension size and the sample size is much larger compared to the other, or when both sizes are comparable. The thesis has the following objectives:

(i) To study large-deviation asymptotics of condition numbers defined as ratios of maximal and minimal eigenvalues of sample covariance matrices and $\beta-$Laguerre ensembles (see Paper I and III);

(ii) To study individual maximal and minimal eigenvalues of sample covariance matrices and $\beta-$Laguerre ensembles in terms of large deviations (see Paper II and IV).

1.3 Motivations

Eigenvalues of a random matrix play a role in many different areas which motivates this research work. Our first motivation is the role of condition numbers as the ratios of maximal and minimal eigenvalues in numerical algebra and in probability theory [50] and [43]. Another motivation is that there is a connection between condition numbers and the so-called *first anti-eigenvalues* defined as $2(\lambda_{max}\lambda_{min})^{1/2}/(\lambda_{max}+\lambda_{min})$ which can be found in many different areas of application (see [25]), where λ_{max} and λ_{min} denote maximal and minimal eigenvalues of a symmetric random matrix. One more motivation is associated to hypothesis testing on whether or not a covariance matrix is a scalar multiple of identity in multivariate statistics [47, Section 7.4], and this can be done based on the maximal eigenvalue of a symmetric random matrix.

In terms of technical motivation, we note that for a $p \times n$ random matrix, the existing large deviations of extremal eigenvalues (of the random matrix multiplied by its transpose) in the literature are based on some restrictive assumptions for instance, $p = o(n/\ln\ln n)$. Hence, the asymptotic results obtained in Paper II are under the condition $p = p(n) = o(n)$ and Paper IV under the condition $p/n \to \kappa \in (0, 1)$. Since large deviations of condition numbers are mostly missing in the literature, we provide results of large deviations of condition numbers in Paper I and III when $p = p(n) = o(n)$ and $p = o(n/\ln n)$ respectively. Our results are inspired by the work of Bianchi et al. [5], Fey et al. [20], and Jiang and Li [29].

1.4 Outline of the thesis

This thesis work has two parts. The first part of the work consists of five chapters and the purpose of each chapter is to facilitate further reading of the papers in part two. Chapter 1 contains a general introduction of the thesis with aims, motivations and outline of the thesis. Chapter 2 reviews some basic concepts concerning the two types of sequences of random matrices considered in this thesis as well as in the papers. Chapter 3 provides theoretical background on large deviations with examples. In Chapter 4, some known results on spectral properties of the two types of sequences of random matrices are presented, together with the contribution of this thesis. Finally, a summary of the appended papers as part of the thesis and plan for future research are given in Chapter 5 of the thesis. The second part of the thesis work contains four papers.

2 Concepts connected to random matrices with independent entries

In this chapter, we introduce two types of sequences of random matrices with independent entries. We start with a definition of a sub-Gaussian random variable, and then introduce the sequence of sample covariance matrices with independent and identically distributed (or i.i.d. in short) sub-Gaussian entries. This sequence contains the well-known real Wishart matrices. To go beyond real Wishart matrices, we introduce another sequence of random matrices called the sequence of $\beta-$Laguerre ensembles, which includes real, complex and quaternion Wishart matrices as special cases.

Definition 2.1 (Sub-Gaussian random variable)**.** ([55, Lemma 5.5])
A random variable X is said to be *sub-Gaussian* if it satisfies one of the following three equivalent properties with parameters $K_i, 1 \leqslant i \leqslant 3$ differing from each other by at most an absolute constant factor:
(i) Tails: $P(|X| > t) \leqslant \exp\{1 - t^2/K_1^2\}$ for all $t \geqslant 0$;
(ii) Moments: $(E|X|^p)^{1/p} \leqslant K_2\sqrt{p}$ for all $p \geqslant 1$;
(iii) Super-exponential moment: $E\exp\{X^2/K_3^2\} \leqslant e$.

If the mean of X is 0, then the above three properties are equivalent to the following property:
(iv) Moment generating function: $E\exp\{tX\} \leqslant \exp\{t^2K_4^2\}$ for all $t \in \mathbb{R}$, and some constant K_4.

The *sub-Gaussian norm* of X is then defined as $\sup_{p\geqslant 1} p^{-1/2}(E|X|^p)^{1/p}$, namely the smallest K_2 in (ii).

2.1 Sample covariance matrices

Sample covariance matrices play an important role in multivariate analysis. For instance, they can be used to check the form of correlation in the data. A sample covariance matrix is commonly defined as follows:

$$\mathbf{S} = (n+1)^{-1} \sum_{j=1}^{n+1} (y_j - \bar{y})(y_j - \bar{y})^T,$$

with $\bar{y} = (n+1)^{-1} \sum_{j=1}^{n+1} y_j$ known as sample mean, where $\{y_j\}_{j=1}^{n+1}$ is a sample from a $p-$variate population $y \sim \mathcal{N}_p(\mu, \mathbf{I}_p)$.

For a $p \times n$ random matrix $\mathbf{X}$ whose entries $X_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n$ are i.i.d. standard normal, if we define $p \times p$ square random matrix $\mathbf{XX}^T$, in the literature it is known that $(n+1)\mathbf{S}$ and $\mathbf{XX}^T$ have the same distribution as shown in [23]. In this case, the two integers p and n can be understood as the sample and the dimension sizes. In this thesis, we consider more general random matrices $\mathbf{X}$ with i.i.d. sub-Gaussian entries. Paper I provides large deviations of condition numbers of $\mathbf{XX}^T$ for fixed p or $p = p(n) = o(n)$, as $n \to \infty$, while Paper II studies the large deviations of individual maximal and minimal eigenvalues of $\mathbf{XX}^T$.

To study the spectral properties of $\mathbf{XX}^T$, in this thesis, without loss of generality, we assume that $p \leqslant n$. If $p > n$ one can study the spectral properties of $\mathbf{X}^T\mathbf{X}$ since the non-trivial eigenvalues of both square matrices $\mathbf{XX}^T$ and $\mathbf{X}^T\mathbf{X}$ are the same, which is explained in Section 2.2.

2.2 Some properties of condition numbers

In this section we lay out some basic properties regarding condition numbers of random matrices. Let $\mathbf{A}$ be $m \times n$ real random matrix and without loss of generality let us assume that $m \leqslant n$. Then $\mathbf{A}$ has m singular values, which are the square roots of the m eigenvalues of $\mathbf{AA}^T$. Let $\sigma = (\sigma_1, \dots, \sigma_m)$ denote the singular values of $\mathbf{A}$ with ordered ones written as $\sigma_{(1)} \geqslant \sigma_{(2)} \geqslant \dots \geqslant \sigma_{(m)}$. The 2-norm condition number $\kappa(\mathbf{A})$ of the matrix $\mathbf{A}$ is defined as

$$\kappa(\mathbf{A}) = \sigma_{max}/\sigma_{min}, \tag{2.1}$$

where $\sigma_{(1)} = \sigma_{max}$ and $\sigma_{(m)} = \sigma_{min}$ are the maximal and minimal singular values of $\mathbf{A}$. The name 2-norm comes from the fact that the maximal singular value coincides with norm of the matrix $\mathbf{A}$, i.e., $||\mathbf{A}||_2 = \sup\{||\mathbf{A}a||_2 : a \in \mathbb{R}^n \text{ with } ||a||_2 = 1\}$, where the 2-norm of a vector $a \in \mathbb{R}^n$ is the Euclidean norm defined as $||a||_2 = (\sum_{i=1}^n a_i^2)^{1/2}$.

Lemma 2.2. *Consider $m \times m$ square matrix $\mathbf{AA}^T$ and $n \times n$ square matrix $\mathbf{A}^T\mathbf{A}$ with $(\lambda_{(1)} \geqslant \dots \geqslant \lambda_{(m)})$ and $(\tilde{\lambda}_{(1)} \geqslant \dots \geqslant \tilde{\lambda}_{(m)}, 0, 0, \dots, 0)$ to denote the ordered eigenvalues of $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$, respectively. Then, the two matrices $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$ have the same non-trivial eigenvalues, namely $\lambda_{(i)} = \tilde{\lambda}_{(i)}$, $1 \leqslant i \leqslant m$.*

Proof. For completeness, a short proof is given here. We show that $\lambda = \lambda_{(i)} = \tilde{\lambda}_{(i)}$, $1 \leqslant i \leqslant m$. Suppose that $\lambda = \lambda_{(i)}$, $1 \leqslant i \leqslant m$ are the eigenvalues of $\mathbf{AA}^T$. We have $\mathbf{AA}^T u = \lambda u$, for some non-zero vector $u \in \mathbb{R}^m$. Then,

$$\begin{aligned} \mathbf{A}^T\mathbf{A}\mathbf{A}^T u &= \mathbf{A}^T \lambda u \qquad \text{(if we multiply by } \mathbf{A}^T \text{ on both sides)} \\ \mathbf{A}^T\mathbf{A}(\mathbf{A}^T u) &= \lambda(\mathbf{A}^T u) \\ \mathbf{A}^T\mathbf{A}v &= \lambda v, \end{aligned}$$

for non-zero vector $v = \mathbf{A}^T u$ and $v \in \mathbb{R}^n$, one can see that λ is also eigenvalue of $\mathbf{A}^T\mathbf{A}$. □

Lemma 2.3. *Let $\kappa(\mathbf{A})$ and $\kappa(\mathbf{A}\mathbf{A}^T)$ denote the condition numbers of the matrix $\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$, respectively. Then, $\kappa^2(\mathbf{A}) = \kappa(\mathbf{A}\mathbf{A}^T)$.*

Proof. The proof is as follows:

$$\begin{aligned} \kappa(\mathbf{A}\mathbf{A}^T) &= \left(\frac{\lambda_{max}(\mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^T)}{\lambda_{min}(\mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^T)}\right)^{1/2} = \left(\frac{\lambda_{max}(\mathbf{A}\mathbf{A}^T\mathbf{A}\mathbf{X}^T)}{\lambda_{min}(\mathbf{A}\mathbf{A}^T\mathbf{A}\mathbf{A}^T)}\right)^{1/2} \\ &= \left(\frac{\lambda^2_{max}(\mathbf{A}\mathbf{A}^T)}{\lambda^2_{min}(\mathbf{A}\mathbf{A}^T)}\right)^{1/2} = \frac{\lambda_{max}(\mathbf{A}\mathbf{A}^T)}{\lambda_{min}(\mathbf{X}\mathbf{A}^T)} = \kappa^2(\mathbf{A}). \end{aligned}$$

□

Remark 1. Note that the matrix $\mathbf{A}^T\mathbf{A}$ is singular when $m < n$, for which the condition number $\kappa(\mathbf{A}^T\mathbf{A}) = \infty$.

In the next section, we define the other type of sequence of random matrices considered in this thesis.

2.3 $\beta-$Laguerre ensembles

In random matrix theory, the size of a matrix and its distribution from which the entries have been chosen are important in the study of the behaviors of eigenvalues of the random matrix. When a $p \times n$ random matrix $\mathbf{X}$ is constructed from a standard normal distribution, the entries can either be i.i.d. real, complex or quaternion standard normal random variables. The corresponding $\mathbf{X}\mathbf{X}^T$ is known as $\beta-$Wishart ensembles (or $\beta-$Laguerre ensembles as technical name) with $\beta = 1, 2, 4$ to point out that when $\beta = 1$ the entries of random matrix $\mathbf{X}$ are real, when $\beta = 2$ the entries are complex, and when $\beta = 4$ the entries are quaternion standard normal random variables. The notion ^{T} stands for transpose when $\beta = 1$, complex conjugate transpose when $\beta = 2$ and dual transpose when $\beta = 4$.

There exist explicit joint density functions of all eigenvalues of these classical $\beta-$Laguerre ensembles. Let $\lambda = (\lambda_1, \ldots, \lambda_p)$ be the eigenvalues of $\mathbf{X}\mathbf{X}^T$, then the joint probability density function of λ is (see [27]),

$$f(\lambda; p, n, \beta) = c(p, n, \beta) \cdot \exp\left\{-\sum_{i=1}^{p} \lambda_i/2\right\} \cdot \prod_{1 \leqslant i < j \leqslant p} |\lambda_i - \lambda_j|^\beta \cdot \prod_{i=1}^{p} \lambda_i^{\beta(n-p+1)/2-1}, \tag{2.2}$$

where $c(p, n, \beta) = 2^{-\beta np/2} \prod_{j=1}^{p} \Gamma(1 + \beta/2) \left[\Gamma(1 + \beta j/2)\Gamma(\beta(n - p + j)/2)\right]^{-1}$, and $\beta = 1, 2, 4$. It had been actually believed that no other choice of $\beta > 0$ (besides $\beta = 1, 2$ and 4) would correspond to a matrix model $\mathbf{X}_\beta\mathbf{X}_\beta^T$ which can be constructed with entries from a classical distribution until the work [14].

To see this, let define a random matrix $\mathbf{X}_\beta$ of bidiagonal form as follows:

$$\mathbf{X}_\beta \sim \begin{bmatrix} \chi_{\beta n} & & & \\ \chi_{\beta(p-1)} & \chi_{\beta(n-1)} & & \\ & \ddots & \ddots & \\ & & \chi_\beta & \chi_{\beta(n-p+1)} \end{bmatrix}.$$

For $\beta > 0$, the $\beta-$Laguerre ensembles are defined as tridiagonal matrices $\mathbf{X}_\beta\mathbf{X}_\beta^T$ with independent χ distribution entries in a bidiagonal matrix $\mathbf{X}_\beta$. It turns out that, the eigenvalues of the real tridiagonal random matrix models $\mathbf{X}_\beta\mathbf{X}_\beta^T$ with $\beta > 0$ have the same expression of joint density function defined in (2.2) as proven in [14]. In this thesis, we study limiting behaviors of the eigenvalues of $\mathbf{X}_\beta\mathbf{X}_\beta^T$ when n, and possibly p as well, tend to infinity.

More specifically, Paper I provides results about condition numbers of real Wishart matrices ($\beta = 1$) when p is either fixed or $p = p(n) \to \infty$, as $n \to \infty$ with $p = o(n)$. For general $\beta > 0$, Paper III studies the ratios of extremal eigenvalues of $\beta-$Laguerre ensembles for fixed p or $p = p(n) \to \infty$ as $n \to \infty$ with $p = o(n/\ln n)$. In Paper IV large deviations are studied for individual λ_{max} and λ_{min}, and for joint $(\lambda_{max}, \lambda_{min})$ eigenvalues of $\beta-$Laguerre ensembles with general $\beta > 0$ under the assumption $p/n \to \kappa \in (0, 1)$ as $p \to \infty$ and $n \to \infty$. In Chapter 4, we review some existing results in the literature about eigenvalues of $\beta-$Laguerre ensembles.

3 Theoretical background on large deviations

The objective of this chapter is to give a short introduction to the theory of large deviations, which includes the tools to be used in the papers in the second part of the thesis. In this thesis, we focus on some basic results on large deviations such as Cramér's theorem, Gärtner-Ellis theorem and the contraction principle, and how these results appear and are applied in the study of eigenvalues of random matrices in the second part of the thesis. For further reading on large deviations, we refer to [10, 37].

The theory of large deviations is often referred as a theory which studies probabilities of rare events or the events which occur with small frequencies. A fundamental result in the theory of large deviations is Cramér's theorem for empirical means of i.i.d. random variables, which was introduced by H. Cramér in 1928 [8]. Here we recall the theorem and briefly mention how it is applied to study large deviations of extremal eigenvalues and condition numbers of random matrices in Papers I and II.

3.1 Cramér's theorem for i.i.d. random variables

We begin this section with a definition of the large deviation principle (LDP) for a sequence of probability measures.

Definition 3.1 (LDP)**.** ([10, Section 1.2]) Let $\{P_n\}_{n\geqslant 1}$ be a sequence of probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, and $I : \mathbb{R} \to [0, \infty]$ be a lower semi-continuous function. The function I is said to be a good rate function if the set $\{c \in \mathbb{R} : I(c) \leqslant \epsilon\}$ is compact for each $\epsilon \in [0, \infty)$; and the sequence $\{P_n\}_{n\geqslant 1}$ is said to satisfy an LDP, as $n \to \infty$, with a good rate function I and speed n if

(i) for any closed set $F \subset \mathbb{R}$, we have an upper bound:

$$\limsup_{n\to\infty} n^{-1} \ln P_n(F) \leqslant -\inf_{c\in F} I(c); \tag{3.1}$$

(ii) for any open set $G \subset \mathbb{R}$, we have a lower bound:

$$\liminf_{n\to\infty} n^{-1} \ln P_n(G) \geqslant -\inf_{c\in G} I(c). \tag{3.2}$$

Now for a sequence of (real) i.i.d. random variables $\{X_n\}_{n\geqslant 1}$, let P_n denote the distributions of the empirical means $S_n = (X_1 + X_2 + \ldots + X_n)/n$. The Cramér's theorem handles an LDP associated with this family $\{P_n\}_{n\geqslant 1}$ of probability measures.

Theorem 3.2 (Cramér's theorem in $\mathbb{R}$)**.** *[10, Theorem2.2.3] The family $\{P_n\}_{n\geqslant 1}$ satisfies an LDP with a speed n and a rate function $I(c)$ defined as Fenchel-Legendre transform of the logarithmic moment generating function*

$$I(c) = \sup_{\theta\in\mathbb{R}}[\theta \cdot c - \Lambda(\theta)], \quad \forall \ c \in \mathbb{R}, \tag{3.3}$$

where $\Lambda(\theta) = \ln M(\theta)$ and $M(\theta) = E\exp\{\theta X_1\}$.

For an appropriately chosen closed set F and an open set G in the LDP of Cramér's theorem, one can obtain a special case as follows.

Corollary 3.3. *[37, Theorem 1.1] For (real) i.i.d. random variables $\{X_n\}_{n\geqslant 1}$ with finite mean μ and variance σ^2, it holds that,*

$$\lim_{n\to\infty} n^{-1} \ln P\left(S_n \geqslant c\right) = -I(c), \quad \text{for } c > \mu. \tag{3.4}$$

To see how Cramér's theorem is applied in Papers I and II when studying asymptotic lower and upper bounds of condition numbers and extremal eigenvalues of random matrices $\mathbf{XX}^T$, whose entries $X_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n$, are i.i.d. sub-Gaussian random variables, we relate probabilities of condition numbers and extremal eigenvalues to the ones involving n i.i.d. random variables. To see this more explicitly, one can show that (see Papers I and II for more details)

$$P\left(\lambda_{max} \geqslant c\right) = P\left(\exists x \in S : \sum_{i=1}^{n} S_{x,i}^2/n \geqslant c\right),$$

$$P\left(\lambda_{\min} \leqslant c\right) = P\left(\exists x \in S : \sum_{i=1}^{n} S_{x,i}^2/n \leqslant c\right),$$

$$P\left(\lambda_{max}/\lambda_{min} \geqslant c\right) = P\left(\exists \ x, y \in S : x \cdot y = 0 \text{ and } \sum_{i=1}^{n}(S_{x,i}^2 - cS_{y,i}^2)/n \geqslant 0\right),$$

where $S := \{x, y \in \mathbb{R}^p : ||x|| = ||y|| = 1\}$ denotes the unit sphere and $S_{x,i} = \sum_{k=1}^{p} x_k X_{ki}$. Then $\{S_{x,n}^2\}_{n\geqslant 1}$ (and also for $\{(S_{x,n}^2 - S_{y,n}^2)\}_{n\geqslant 1}$) is a sequence of (real) i.i.d. random variables, for which Cramér's theorem can be employed.

Here, we list several explicit rate functions for i.i.d. random variables drawn from known distributions.

Example 3.4 (Rate functions). Suppose that $\{X_n\}_{n\geqslant 1}$ are i.i.d. random variables.

(i) If $X_1 \sim \mathcal{N}(\mu, \sigma^2)$, then $\Lambda(\theta) = \ln E\exp\{\theta X_1\} = \mu\theta + (\sigma^2/2)\theta^2$, which gives $I(c) = \sup_{\theta\in\mathbb{R}}[\theta \cdot c - \Lambda(\theta)] = (c-\mu)^2/(2\sigma^2)$. Hence, for $c > \mu$,

$$\lim_{n\to\infty} n^{-1} \ln P(S_n \geqslant c) = -\frac{(c-\mu)^2}{2\sigma^2}.$$

(ii) If X_1 is Bernoulli random variable with $P(X_1 = 1) = P(X_1 = -1) = 1/2$, then $\Lambda(\theta) = \ln E\exp\{\theta X_1\} = \ln((\exp\{\theta\} + \exp\{-\theta\})/2)$, and $I(c) = 1/2(1+c)\ln(1+c) + 1/2(1-c)\ln(1-c)$ for $c \in [-1, +1]$, and ∞ otherwise. Therefore, for $c > 0$

$$\lim_{n\to\infty} n^{-1} \ln P(S_n \geqslant c) = -I(c).$$

Note that Cramér's theorem deals with i.i.d. random variables. When it comes to random variables which are not i.i.d., Gärtner-Ellis theorem is often used.

3.2 Gärtner-Ellis theorem

In Papers I and III, in the case for which Cramér's theorem can not be applied immediately to study large-deviation asymptotics of condition numbers of sample covariance matrices with i.i.d. sub-Gaussian random variables and $\beta-$Laguerre ensembles, we employ the Gärtner-Ellis theorem. To this end, let us recall a definition.

Definition 3.5 (Exposed point and essential smoothness)**.** [10, Section 2.3]
(a) A point $c \in \mathbb{R}$ is said to be an exposed point of a function I if there exists a point $a \in \mathbb{R}$ such that $a \cdot (c-d) < I(c) - I(d)$ for $\forall\ c \neq d$. Such a is called an exposing hyperplane.

(b) A convex function Λ is said to be essentially smooth if Λ is a differentiable function on the nonempty interior D^o_Λ of $D_\Lambda := \{c \in \mathbb{R} : \Lambda(c) < \infty\}$ of Λ, and $|\Lambda'(c_n)| \to \infty$ when $c_n \to c$, for c on the boundary of D^o_Λ.

Theorem 3.6 (Gärtner-Ellis theorem)**.** *[10, Section 2.3]*
Suppose that $\{X_n\}_{n\geqslant 1}$ are defined on a probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$ such that the limit exists: $\Lambda(\theta) = \lim_{n\to\infty} n^{-1} \ln E\exp\{n\theta X_n\}$ for $\theta \in \mathbb{R}$ and $0 \in D^o_\Lambda$. Let $I(c)$ be given by the Fenchel-Legendre transform (3.3) *and $\mathcal{F}$ denote the set of exposed points of I whose exposing hyperplane is in D^o_Λ. Then,*

(i) for any closed set $F \subset \mathbb{R}$, we have upper bound:

$$\limsup_{n\to\infty} n^{-1} \ln P(X_n \in F) \leqslant -\inf_{c\in F} I(c); \tag{3.5}$$

(ii) for any open set $G \subset \mathbb{R}$, we have lower bound:

$$\liminf_{n\to\infty} n^{-1} \ln P(X_n \in G) \geqslant -\inf_{c\in G\cap\mathcal{F}} I(c). \tag{3.6}$$

In addition, if Λ is essentially smooth and lower semi-continuous, then $\{X_n\}_{n\geqslant 1}$ satisfies an LDP with a good rate function $I(c)$.

It can be directly seen that Cramér's theorem for empirical means of i.i.d. random variables is a special case of the Gärtner-Ellis theorem. In Papers I and III, the Gärtner-Ellis theorem was employed to study large deviations of condition numbers of the two types of sequences of random matrices. To explain this more specifically, in Papers I and III, we analyse directly the joint distribution function of the eigenvalues, and delete (by estimating) almost all information except for the maximal and minimal eigenvalues, which leads to ratios of independent summations of Gamma random variables $U_{1,m}$ and $U_{2,m}$ with two different parameters $f(m)$ and $g(m)$ both tending to infinity. One can then rewrite

$$P(U_{1,m}/U_{2,m} \geqslant c) = P\left(\sum_{i=1}^{f(m)} \xi_i - c\sum_{i=1}^{g(m)} \eta_i \geqslant 0\right),$$

with i.i.d. Gamma random variables ξ_i and η_i. If $f(m) = g(m)$, then Cramér's theorem applies as $\{U_{1,m}/U_{2,m} \geqslant c\} = \frac{1}{f(m)} \sum_{i=1}^{f(m)} (\xi_i - c\eta_i)$. However, when $f(m)$ and $g(m)$ are different, in this case Cramér's theorem is not applicable, instead the Gärtner-Ellis theorem works well. That is, we define

$$Z_m = \frac{1}{f(m)}\left(\sum_{i=1}^{f(m)} \xi_i - c\sum_{i=1}^{g(m)} \eta_i\right).$$

Then, the logarithmic moment generating function of Z_m satisfies the conditions of the Gärtner-Ellis theorem.

3.3 Contraction principle

In order to examine the relation between large deviations of individual extremal eigenvalues λ_{max} and λ_{min} and large deviations of the joint $(\lambda_{max}, \lambda_{min})$, let us recall the contraction principle.

Definition 3.7 (Contraction principle). [10, Section 4.2] Let $\{X_n\}_{n\geqslant 1}$ be a sequence of random variables satisfying an LDP with a good rate function I and f be continuous function, then $\{f(X_n)\}_{n\geqslant 1}$ satisfies an LDP with a good rate function

$$J(y) := \inf_x\{I(x) : f(x) = y\}.$$

To explore possible connections between large deviations of the ratios of extremal eigenvalues $\lambda_{max}/\lambda_{min}$ and large deviations of joint extremal eigenvalues $(\lambda_{max}, \lambda_{min})$, let us first state a result that can be proved in the spirit of the contraction principle.

Lemma 3.8. *[10, Excercise 4.2.7])*
Let $\{X_n\}_{n\geqslant 1}$ and $\{Y_n\}_{n\geqslant 1}$ be two independent sequences satisfying LDPs with good rate functions $I_1(x)$ and $I_2(y)$, respectively, and both sequences are exponentially tight. Then the sequence $\{f(X_n, Y_n)\}_{n\geqslant 1}$ with continuous f satisfies an LDP with a good rate function

$$J(z) = \inf_{x,y}\{I_1(x) + I_2(y) : f(x,y) = z\}.$$

If large deviations for individual λ_{max} and λ_{min} are available, then large deviations for the joint $(\lambda_{max}, \lambda_{min})$ would follow from the contraction principle if λ_{max} and λ_{min} were independent (which is not the case) and exponentially tight. Large deviations for the joint $(\lambda_{max}, \lambda_{min})$ is desirable since it would imply large deviations of the condition numbers $\lambda_{max}/\lambda_{min}$, which follows directly from the contraction principle with a continuous function $f(\lambda_{max}, \lambda_{min}) := \lambda_{max}/\lambda_{min}$, as $P(\lambda_{min} > 0) = 1$. In Paper IV we establish some results of large deviations for both the individual and joint extremal eigenvalues of $\beta-$Laguerre ensembles.

4 Spectral properties of random matrices

In this chapter we review some known spectral results from the literature for the two types of sequences of random matrices considered in the thesis. Two cases are considered separately: (i) the bulk case which deals with all the eigenvalues, and (ii) the extremal case which is the extremal (maximal and minimal) eigenvalues. The connections between our contributions (in the thesis) and known results (in the literature) are also mentioned. It is remarked here that the thesis does not make any contributions at all to the bulk case. However, for the sake of completeness here we still briefly review the known results.

4.1 The bulk case

For each type of sequence of random matrices considered in Chapter 2 (namely, the sequence of sample covariance matrices $\mathbf{X}\mathbf{X}^T$ or the sequence of $\beta-$Laguerre ensembles $\mathbf{X}_\beta\mathbf{X}_\beta^T$), if we consider the empirical distributions of all the eigenvalues, then there are basically two types of convergence results, depending on the relation of p and n. When $p/n \to \kappa \in (0,1]$, the (scaled) empirical distributions converge to the so-called Marčenko-Pastur distribution μ_{MP} given as

$$\mu_{MP}(dx) = 1_{(\kappa_-,\kappa_+)}(x) \cdot \frac{\sqrt{(\kappa_+ - x)(x-\kappa_-)}}{2\pi\kappa x} dx, \tag{4.1}$$

where $\kappa_\pm = (1 \pm \kappa^{1/2})^2$; see [38, Theorem 3.2]) for the sequence of sample covariance matrices, and [13, Theorem 6.5.1] for the sequence of $\beta-$Laguerre ensembles. When $p(n)/n \to 0$, the empirical distributions after proper scaling usually converge to the semi-circle distribution; see [2] for the sequence of sample covariance matrices, and [29, Preposition 3] for the sequence of $\beta-$Laguerre ensembles.

In addition to the aforementioned convergence, in the literature there are also results on large deviations of the empirical distributions deviating from the limiting distributions under various restrictions on p and n; see for example [40] for the sequence of sample covariance matrices, and [29, Theorem 4] for the sequence of $\beta-$Laguerre ensembles.

Next, we move to the limiting behaviors of the extremal eigenvalues of the two types of sequences of random matrices considered in this thesis.

4.2 The extremal case

We start with weak convergence of the maximal and minimal eigenvalues of the two types of sequences of random matrices considered in Chapter 2. In 1996 the distribution of the maximal eigenvalue of a Gaussian hermitian or symmetric matrix was first studied by Tracy and Widom [49]. Later the results were extended to matrices of the form $\mathbf{X}_\beta \mathbf{X}_\beta^T$ where $\mathbf{X}_\beta$ are $p \times n$ matrices with $\beta = 1, 2, 4$ (i.e., entries are real, complex or quaternion) for both p and n being large. When $\beta = 1$ or 2, limiting distributions of the maximal eigenvalue of $\mathbf{X}_\beta \mathbf{X}_\beta^T$ were established in [30] (for $\beta = 2$) and [31] (for $\beta = 1$). These results show that the maximal eigenvalue weakly converges to the $\beta-$Tracy-Widom law. In the same way, the minimal eigenvalue also weakly converges to the $\beta-$Tracy-Widom law (see [4] for $\beta = 2$). The most complete results of this type seem to be the ones in [41], which are recalled in the following lemma.

Lemma 4.1 (Tracy-Widom law of extremal eigenvalues). *[41] Let $\mathbf{X}_\beta \mathbf{X}_\beta^T$ be the family of $\beta-$Laguerre ensembles defined in Chapter 2, and λ_{max} and λ_{min} be the maximal and minimal eigenvalues. Then, for $p(n)/n \to \kappa$ with $p = p(n) \to \infty$, as $n \to \infty$,*

(i) if $\kappa \in [0, 1]$:

$$(\sqrt{pn})^{1/3} \left(\lambda_{max}/\beta - (\sqrt{n} + \sqrt{p})^2 \right) / (\sqrt{n} + \sqrt{p})^{4/3} \to^d TW_\beta, \tag{4.2}$$

(ii) if $\kappa \in [0, 1)$:

$$(\sqrt{pn})^{1/3} \left((\sqrt{n} - \sqrt{p})^2 - \lambda_{min}/\beta \right) / (\sqrt{n} - \sqrt{p})^{4/3} \to^d TW_\beta, \tag{4.3}$$

where TW_β is the general β Tracy-Widom law.

The above results imply the following laws of large numbers (see also [29]): with $p = p(n) \to \infty$, as $n \to \infty$ and $p(n)/n \to \kappa$,

$$\lambda_{max}/n \to (1 + \kappa^{1/2})^2 \beta, \quad \kappa \in [0, 1], \quad \text{in probability;} \tag{4.4}$$

and

$$\lambda_{min}/n \to (1 - \kappa^{1/2})^2 \beta, \quad \kappa \in [0, 1), \quad \text{in probability.} \tag{4.5}$$

When one considers the sequence of sample covariance matrices $\mathbf{X}\mathbf{X}^T$ with sub-Gaussian entries, quite similar results as (4.4) and (4.5) also hold under suitable assumptions for the corresponding maximal and minimal eigenvalues; see [61] and [3].

Large deviations for individual extremal eigenvalues

With the laws of large numbers in (4.4) and (4.5), it is natural to study large deviations of individual λ_{max} and λ_{min}. Indeed, in [29] for the sequence of $\beta-$Laguerre ensembles (with a general $\beta > 0$), large deviations of individual λ_{max} and λ_{min} have been well studied mostly under the assumption $p/n \to 0$. But what if p and n are comparable (namely $p/n \to \kappa \in (0,1)$)? It turns out that the various estimates used in [29] and the references therein are not precise enough, and other tools are needed for the study of large deviations. In Paper IV, we employ Laplace's method to establish large deviations of individual λ_{max} and λ_{min} (and joint $(\lambda_{max}, \lambda_{min})$) for all $\beta > 0$, under the assumption $p/n \to \kappa \in (0,1)$ as $n \to \infty$ and $p \to \infty$ at the same time, which generalizes the special case $\beta = 2$ considered in [5]. The main results of Paper IV are recalled as follows in Theorems 4.2 and 4.3.

Theorem 4.2. *(i) The sequence $\{\lambda_{max}/(2n)\}$ satisfies an LDP with a speed n and a suitably defined rate function $I(x)$:*

$$\limsup_{n\to\infty} n^{-1} \ln P\left(\lambda_{max}/(2n) \in F\right) \leqslant -\inf_{x\in F} I(x), \quad \text{for any closed set } F, \tag{4.6}$$

$$\liminf_{n\to\infty} n^{-1} \ln P\left(\lambda_{max}/(2n) \in O\right) \geqslant -\inf_{x\in O} I(x), \quad \text{for any open set } O. \tag{4.7}$$

(ii) The sequence $\{\lambda_{min}/(2n)\}$ satisfies an LDP with a speed n and a suitably defined rate function $J(y)$:

$$\limsup_{n\to\infty} n^{-1} \ln P\left(\lambda_{min}/(2n) \in F\right) \leqslant -\inf_{y\in F} J(y), \quad \text{for any closed set } F, \tag{4.8}$$

$$\liminf_{n\to\infty} n^{-1} \ln P\left(\lambda_{min}/(2n) \in O\right) \geqslant -\inf_{y\in O} J(y), \quad \text{for any open set } O. \tag{4.9}$$

Theorem 4.3. *For any $x_0 \geqslant (1+\kappa^{1/2})^2\beta/2$ and $y_0 \leqslant (1-\kappa^{1/2})^2\beta/2$, it holds that*

$$\lim_{n\to\infty} n^{-1} \ln P\left(\lambda_{max}/(2n) \geqslant x_0, \lambda_{min}/(2n) \leqslant y_0\right) = -I(x_0) - J(y_0). \tag{4.10}$$

For the sequence of sample covariance matrices $\mathbf{X}\mathbf{X}^T$, the authors in [20] studied large deviations of individual λ_{max} and λ_{min} under two main restrictions: (i) the i.i.d. entries are either (real) standard normal or symmetric and bounded; (ii) the dimension size p and the sample size n satisfy $p = o(n/\ln\ln n)$ as $n \to \infty$. In Paper II, based on contraction inequalities, we improved the results by (i) allowing the i.i.d. entries to be sub-Gaussian, and (ii) imposing the relaxed assumption $p = o(n)$. The main result of Paper II is recalled in the following Theorem 4.4.

Theorem 4.4. *Supposed that the entries $\{X_{ij}\}_{1\leqslant i\leqslant p, 1\leqslant j\leqslant n}$ of $\mathbf{X}$ are i.i.d. sub-Gaussian with zero mean and unit variance. Then, for $p = p(n) \to \infty$ with $p(n) = o(n)$ as $n \to \infty$, and a suitably defined rate function I :*

(i) for any $c \geqslant 1$ it holds that

$$\liminf_{n\to\infty} n^{-1} \ln P\left(\lambda_{\max} \geqslant c\right) \geqslant -I(c), \tag{4.11}$$

$$\limsup_{n\to\infty} n^{-1} \ln P\left(\lambda_{\max} \geqslant c\right) \leqslant -\lim_{\epsilon\to 0} I(c-\epsilon); \tag{4.12}$$

(ii) for any $0 \leqslant c \leqslant 1$ it holds that,

$$\liminf_{n\to\infty} n^{-1} \ln P\left(\lambda_{\min} \leqslant c\right) \geqslant -I(c), \tag{4.13}$$

$$\limsup_{n\to\infty} n^{-1} \ln P\left(\lambda_{\min} \leqslant c\right) \leqslant -\lim_{\epsilon\to 0} I(c+\epsilon). \tag{4.14}$$

Large deviations for condition numbers

Large deviations for condition numbers $\kappa(p,n) = (\lambda_{max}/\lambda_{min})^{1/2}$ of the aforementioned two types of sequences of random matrices are mostly unknown. A partial reason can be shown as follows. As discussed already in Section 3.3, large deviations of $\kappa(p,n)$ would follow directly from large deviations of the joint $(\lambda_{max}, \lambda_{min})$ based on the contraction principle. However, given large deviations of individual λ_{max} and λ_{min}, large deviations of the joint $(\lambda_{max}, \lambda_{min})$ won't follow from the contraction principle, as λ_{max} and λ_{min} are not independent. Therefore, in order to establish large deviations of the condition numbers $\kappa(p,n)$, one either establishes large deviations of the joint $(\lambda_{max}, \lambda_{min})$, or investigates the condition numbers directly. In Papers I and III, we directly study the condition numbers $\kappa(p,n)$, and derive large deviations results based on careful analysis of the joint density function of all eigenvalues.

The main result of Paper I for the sequence of sample covariance matrices $\mathbf{X}\mathbf{X}^T$ is recalled in Theorem 4.5.

Theorem 4.5. *Suppose that the entries $X_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n$, are i.i.d. sub-Gaussian with zero mean and unit variance. Then for any $c \geqslant 1$ and suitably defined rate functions $I_{p,0}$ and $I_{\infty,0}$, it holds that for fixed p,*

$$\lim_{n\to\infty} n^{-1} \ln P\left(k^2(p,n) \geqslant c\right) = -I_{p,0}(c), \tag{4.15}$$

and for $p = p(n) \to \infty$ with $p(n) = o(n)$,

$$\lim_{n\to\infty} n^{-1} \ln P\left(k^2(p,n) \geqslant c\right) = -I_{\infty,0}(c). \tag{4.16}$$

Since standard normal distribution is sub-Gaussian, a very special case of Theorem 4.5 is the real central Wishart matrix $W_p(n, n^{-1}\mathbf{I})$ for which the entries $X_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n$, are i.i.d. standard normal $\mathcal{N}(0,1)$.

Corollary 4.6. *Suppose that the entries $X_{ij}, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n$, are i.i.d. standard normal $\mathcal{N}(0,1)$. Then for any $c \geqslant 1$ it holds that*

$$\lim_{n\to\infty} n^{-1} \ln P\left(k^2(p,n) \geqslant c\right) = -2^{-1} \ln\left[(c+1)^2/(4c)\right] \tag{4.17}$$

when p is fixed or $p = p(n) \to \infty$ with $p(n) = o(n)$.

The main result of Paper III for the sequence of β-Laguerre ensembles is recalled in Theorem 4.7.

Theorem 4.7. *For β-Laguerre ensembles whose joint probability density of eigenvalues is given by* (2.2)*, suppose that p is either fixed or $p = p(n) \to \infty$ as $n \to \infty$ with $p = o(n/\ln n)$. Then it holds that, for any $\beta > 0$ and $c \geqslant 1$,*

$$\lim_{n\to\infty} n^{-1} \ln P(\lambda_{\max}/\lambda_{\min} \geqslant c) = -(\beta/2) \ln\left[(c+1)^2/(4c)\right].$$

5 Summary of papers, conclusion and further research

This chapter gives an overview of the results established in part two of this thesis and the plan for future research.

5.1 Summary of papers

Paper I: Large-deviation asymptotics of condition numbers of random matrices

Sub-Gaussian random variables contain random variables such as the normal (or Gaussian), Bernoulli and bounded random variables. In this paper, we study large deviations of condition numbers of random matrices with sub-Gaussian random entries.

Let $\mathbf{X}$ be a $p \times n$ random matrix with zero mean and unit variance i.i.d. sub-Gaussian random entries. For fixed p or $p = p(n) = o(n)$ as $n \to \infty$, we derive large deviations of condition numbers of $\mathbf{W}_{p\times p} := \mathbf{X}\mathbf{X}^T/n$. Note that $\mathbf{W}_{p\times p}$ includes the real central Wishart matrix as a special case.

Paper II: Large deviations of extremal eigenvalues of sample covariance matrices

For a $p \times n$ random matrix $\mathbf{X}$ with i.i.d. sub-Gaussian random entries, the paper studies large deviations of extremal eigenvalues of $\mathbf{X}\mathbf{X}^T/n$ under $p = p(n) \to \infty$ with $p = o(n)$ as $n \to \infty$. This improves an existing result in the literature [20] in which p is more restrictive $p = o(n/\ln\ln n)$.

Paper III: On the ratio of extremal eigenvalues of $\beta-$Laguerre ensembles

Suppose that $\mathbf{X}$ is a $p \times n$ random matrix with i.i.d. entries. When the entries of $\mathbf{X}$ are real ($\beta = 1$), complex ($\beta = 2$) or quaternion ($\beta = 4$) normal random variable, we have three special matrix models of the form $\mathbf{X}\mathbf{X}^T$ under the name of real, complex or quaternion Wishart matrix model. When $\beta > 0$, $\mathbf{X}_\beta \mathbf{X}_\beta^T$ are real tridiagonal random matrix models with independent $\chi-$distribution entries and the models are known as $\beta-$Laguerre ensembles (or $\beta-$Wishart ensembles) in the literature. In this paper we present and prove results of the ratio of extremal eigenvalues of $\beta-$Laguerre ensembles for p either fixed or $p = p(n) \to \infty$ with $p = o(n/\ln n)$ as $n \to \infty$ in terms of large deviations.

Paper IV: Large deviations of extremal eigenvalues of $\beta-$Laguerre ensembles

In this paper, we consider $\beta-$Laguerre (or Wishart) ensembles $\mathbf{X}_\beta \mathbf{X}_\beta^T$ with a general $\beta > 0$. Under the assumption $p/n \to \kappa \in (0, 1)$, we derive large deviations for the extremal (namely, maximal and minimal) eigenvalues. This generalizes a result considered in [5] for $\beta = 2$.

5.2 Further research

Concerning the future research, overall we are planning to apply the results derived in the papers to real life problems in Rwanda to solve certain problems encountered by companies.

In this thesis, we have studied large deviations for the extremal eigenvalues and the condition numbers of the two types of sequences of random matrices: the sequence of sample covariance matrices and the sequence of $\beta-$Laguerre ensembles. On one hand, it would be challenging to improve the conditions imposed on p and n, which will be our future work. For instance, in Paper III, it is not an easy task to improve the condition from $p = o(n/\ln n)$ to $p = o(n)$. On the other hand, it would be interesting to examine whether or not similar results can be derived as well for other types of sequences of random matrices (for example, sequence of $\beta-$Hermite ensembles). It has been remarked in [29] that the weak limit of the eigenvalues of $\beta-$Laguerre ensembles is the one of the eigenvalues of $\beta-$Hermite ensembles (after suitable scaling) when $p = o(n^{1/3})$. It is expected that there are some connections between the corresponding large deviations.

Bibliography

[1] A. L. Andrew. "Choosing test matrices for numerical software". In: *Computational Techniques and Applications Conference: CTAC-89*. 1990, pp. 561–567.

[2] Z. D. Bai and Y. Q. Yin. "Convergence to the semicircle law". In: *The Annals of Probability* 16.2 (1988), pp. 863–875.

[3] Z. D. Bai and Y. Q. Yin. "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix". In: *Advances In Statistics*. World Scientific, 2008, pp. 108–127.

[4] T. H. Baker, P. J. Forrester, and P. A. Pearce. "Random matrix ensembles with an effective extensive external charge". In: *Journal of Physics A: Mathematical and General* 31.29 (1998), p. 6087.

[5] P. Bianchi, M. Debbah, M. Maida, and J. Najim. "Performance of statistical tests for single-source detection using random matrix theory". In: *IEEE Transactions on Information Theory* 57.4 (2011), pp. 2400–2419.

[6] B. Bollobás. "Random graphs". In: *Modern graph theory*. Springer, 1998, pp. 215–252.

[7] A. G. Constantine. "Some non-central distribution problems in multivariate analysis". In: *The Annals of Mathematical Statistics* 34.4 (1963), pp. 1270–1285.

[8] H. Cramér and H. Touchette (translator). "On a new limit theorem in probability theory (Translation of'Sur un nouveau théorème-limite de la théorie des probabilités')". In: *ArXiv e-prints* (2018), arXiv–1802.

[9] P. Deift. "Integrable systems and combinatorial theory". In: *Notices AMS* 47 (2000), pp. 631–640.

[10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer-Verlag, New York. 1998.

[11] J. W. Demmel. "The probability that a numerical analysis problem is difficult". In: *Mathematics of Computation* 50.182 (1988), pp. 449–480.

[12] P. Diaconis. *Group representations in probability and statistics*. Lecture notes-monograph series, 1988.

[13] I. Dumitriu. "Eigenvalue statistics for beta-ensembles". PhD thesis. Massachusetts Institute of Technology, 2003.

[14] I. Dumitriu and A. Edelman. "Matrix models for beta ensembles". In: *Journal of Mathematical Physics* 43.11 (2002), pp. 5830–5847.

[15] F. J. Dyson. "Statistical theory of the energy levels of complex systems. I". In: *Journal of Mathematical Physics* 3.1 (1962), pp. 140–156.

[16] F. J. Dyson. "The threefold way. Algebraic structure of symmetry groups and ensembles in quantum mechanics". In: *Journal of Mathematical Physics* 3.6 (1962), pp. 1199–1215.

[17] A. Edelman. "On the distribution of a scaled condition number". In: *Mathematics of Computation* 58.197 (1992), pp. 185–190.

[18] P. Erdős and A. Rényi. "On the evolution of random graphs". In: *Publ. Math. Inst. Hung. Acad. Sci* 5.1 (1960), pp. 17–60.

[19] I. J. Farkas, I. Derényi, A. L. Barabási, and T. Vicsek. "Spectra of "real-world" graphs: Beyond the semicircle law". In: *Physical Review E* 64.2 (2001), p. 026704.

[20] A. Fey, R. Van Der Hofstad, and M. J. Klok. "Large deviations for eigenvalues of sample covariance matrices, with applications to mobile communication systems". In: *Advances in Applied Probability* 40.4 (2008), pp. 1048–1071.

[21] M. Gaudin. "Sur la loi limite de l'espacement des valeurs propres d'une matrice aléatoire". In: *Nuclear Physics* 25 (1961), pp. 447–458.

[22] I. P. Goulden and D. M. Jackson. "Maps in locally orientable surfaces and integrals over real symmetric surfaces". In: *Canadian Journal of Mathematics* 49.5 (1997), pp. 865–882.

[23] A. Gupta and D. Nagar. *Matrix variate distributions*. Chapman & Hall/CRC, Boca Raton,FL, 1999.

[24] I. I. Gurevich and M. I. Pevsner. "Repulsion of nuclear levels". In: *Nuclear Physics* 2.5 (1956), pp. 575–581.

[25] K. E. Gustafson. *Antieigenvalue Analysis: with Applications to Numerical Analysis, Wavelets, Statistics, Quantum Mechanics, Finance and Optimization*. World Scientific, New Jersey, 2011.

[26] P. J. Hanlon, R. P. Stanley, and J. R. Stembridge. "Some combinatorial aspects of the spectra of normally distributed random matrices". In: *Contemporary Mathematics* 138 (1992), pp. 151–174.

[27] A. T. James. "Distributions of matrix variates and latent roots derived from normal samples". In: *The Annals of Mathematical Statistics* 35.2 (1964), pp. 475–501.

[28] A. T. James. "The distribution of the latent roots of the covariance matrix". In: *The Annals of Mathematical Statistics* (1960), pp. 151–158.

[29] T. Jiang and D. Li. "Approximation of rectangular beta-Laguerre ensembles and large deviations". In: *Journal of Theoretical Probability* 28.3 (2015), pp. 804–847.

[30] K. Johansson. "Shape fluctuations and random matrices". In: *Communications in Mathematical Physics* 209.2 (2000), pp. 437–476.

[31] I. M. Johnstone. "On the distribution of the largest eigenvalue in principal components analysis". In: *The Annals of Statistics* 29.2 (2001), pp. 295–327.

[32] E. Khan. "Random matrices, information theory and physics: new results, new connections". In: *Informational Processes* 5.1 (2005), pp. 87–99.

[33] T. Kollo and D. von Rosen. *Advanced Multivariate Statistics with Matrices*. Springer Science & Business Media, Dordrecht, 2005.

[34] B. D. McKay. "The expected eigenvalue distribution of a large regular graph". In: *Linear Algebra and its Applications* 40 (1981), pp. 203–216.

[35] R. J. Muirhead. *Aspects of Multivariate Statistical Theory*. John Wiley & Sons,New York, 1982.

[36] A. Okounkov. "Random matrices and random permutations". In: *International Mathematics Research Notices* 2000.20 (2000), pp. 1043–1095.

[37] E. Olivieri and M. E. Vares. *Large Deviations and Metastability*. 100. Cambridge University Press, 2005.

[38] D. Paul and A. Aue. "Random matrix theory in statistics: A review". In: *Journal of Statistical Planning and Inference* 150 (2014), pp. 1–29.

[39] K. Pearson. "LIII. On lines and planes of closest fit to systems of points in space". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (1901), pp. 559–572.

[40] D. Petz and F. Hiai. "Logarithmic energy as an entropy functional". In: *Contemporary Mathematics* 217 (1998), pp. 205–221.

[41] J. Ramirez, B. Rider, and B. Virág. "Beta ensembles, stochastic Airy spectrum, and a diffusion". In: *Journal of the American Mathematical Society* 24.4 (2011), pp. 919–944.

[42] S. N. Roy. "On a heuristic method of test construction and its use in multivariate analysis". In: *The Annals of Mathematical Statistics* 24.2 (1953), pp. 220–238.

[43] M. Rudelson. "Invertibility of random matrices: norm of the inverse". In: *Annals of Mathematics* (2008), pp. 575–600.

[44] C. Schensted. "Longest increasing and decreasing subsequences". In: *Canadian Journal of Mathematics* 13 (1961), pp. 179–191.

[45] J. W. Silverstein and P. L. Combettes. "Large dimensional random matrix theory for signal detection and estimation in array processing". In: *SSAP, Workshop on Statistical Signal and Array Processing*. Citeseer. 1992, pp. 276–279.

[46] M. Singull, D. Uwamariya, and X. Yang. "Large-deviation asymptotics of condition numbers of random matrices". In: *Journal of Applied Probability* 58.4 (2021), pp. 1114–1130.

[47] M. S. Srivastava and C. G. Khatri. *An Introduction to Multivariate Statistics*. North-Holland, New York-Oxford, 1979.

[48] E. Telatar. "Capacity of multi-antenna Gaussian channels". In: *European Transactions on Telecommunications* 10.6 (1999), pp. 585–595.

[49] C. A. Tracy and H. Widom. "On orthogonal and symplectic matrix ensembles". In: *Communications in Mathematical Physics* 177.3 (1996), pp. 727–754.

[50] L. N. Trefethen and D. Bau III. *Numerical Linear Algebra*. Vol. 50. Siam, 1997.

[51] D. N. C. Tse. "Multiuser receivers, random matrices and free probability". In: *In 37th Annual Allerton Conference on Communication, Control and Computing*. 1999.

[52] D. Uwamariya and X. Yang. "Large deviations of extremal eigenvalues of $\beta-$Laguerre ensembles". In: *Submitted* (2023).

[53] D. Uwamariya and X. Yang. "Large deviations of extremal eigenvalues of sample covariance matrices". In: *Accepted in Journal of Applied Probability* (2023).

[54] D. Uwamariya and X. Yang. "On the ratio of extremal eigenvalues of $\beta-$Laguerre ensembles". In: *Submitted* (2022).

[55] R. Vershynin. "Introduction to the non-asymptotic analysis of random matrices". In: *arXiv preprint arXiv:1011.3027* (2010).

[56] P. Viswanath, D. N. C. Tse, and V. Anantharam. "Asymptotically optimal water-filling in vector multiple-access channels". In: *IEEE Transactions on Information Theory* 47.1 (2001), pp. 241–267.

[57] E. P. Wigner. *Characteristic vectors of bordered matrices with infinite dimensions, Annals of Mathematics*. 1955.

[58] E. P. Wigner. "On the distribution of the roots of certain symmetric matrices". In: *Annals of Mathematics* (1958), pp. 325–327.

[59] E. P. Wigner. "On the statistical distribution of the widths and spacings of nuclear resonance levels". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 47. 4. Cambridge University Press. 1951, pp. 790–798.

[60] E. P. Wigner. *Statistical Properties of Real Symmetric Matrices with Many Dimensions*. Princeton University, 1957.

[61] Y. Q Yin, Z. D. Bai, and P. R. Krishnaiah. "On the limit of the largest eigenvalue of the large dimensional sample covariance matrix". In: *Probability Theory and Related Fields* 78.4 (1988), pp. 509–521.

[62] L. Zheng and D. N. C. Tse. "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels". In: *IEEE Transactions on information theory* 49.5 (2003), pp. 1073–1096.

Part II

Papers

Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

https://doi.org/10.3384/9789180751575

FACULTY OF SCIENCE AND ENGINEERING

Linköping Studies in Science and Technology, Dissertation No. 2313, 2023
Department of Mathematics

Linköping University
SE-581 83 Linköping, Sweden

www.liu.se

LINKÖPING
UNIVERSITY