Small Deviations for the β-Jacobi Ensemble

by

Benjamin Katz-Moses

B.A., Claremont McKenna College, 2006

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

2012

This thesis entitled:
Small Deviations for the β-Jacobi Ensemble
written by Benjamin Katz-Moses
has been approved for the Department of Mathematics

Brian Rider

Prof. Florian Sobieczky

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Katz-Moses, Benjamin (Ph.D., Mathematics)

Small Deviations for the β-Jacobi Ensemble

Thesis directed by Prof. Brian Rider

Classical random matrix theory has its roots in Mathematical Physics, where the eigenvalues of random matrices with Gaussian entries were used to model the behavior of heavy nuclei. The behavior of random matrix eigenvalues has been observed in a myriad of subjects including combinatorics, quantum mechanics, and statistical mechanics.

Placing a measure on the space of $n \times n$ random matrices with specific entry conditions produces what is called a β-ensemble, which can be described by a joint eigenvalue probability distribution function. Prior to 2002, random matrix models only existed for these ensembles when $\beta = 1, 2$, and 4. In 2002, Dumitriu and Edelman produced tridiagonal matrix models for the β-Hermite and β-Leguerre ensembles for the general parameter $\beta > 0$. This groundbreaking work opened the door to studying β-ensembles for all $\beta > 0$.

Small deviation inequalities describe the rate at which random objects concentrate around a distribution. Optimally, the distribution of the object in question should begin to take on the shape of the limiting distribution. Small deviation inequalities for the Hermite, Leguerre, and Jacobi unitary ensembles ($\beta = 2$) were established by Ledoux in [13]. In [16], Ledoux and Rider showed the Hermite and Leguerre inequalities hold for more general β.

This dissertation establishes various small deviation inequalities for the largest eigenvalue of the β-Jacobi Ensemble when $\beta \geq 1$. Upper bounds for the right and left tails are found be in line with the shape of the Tracy-Widom Distribution. From these small deviation bounds, an upper bound on the variance of the largest eigenvalue is derived. This bound is in accordance with known limit theorems.

Dedication

To my loving parents, Anne Moses and David Katz, who instilled in me the confidence to try and the skills to succeed, and to the memory of my Bubbie, Sophie Katz, whose passion for learning lives on in her family.

Acknowledgements

My deepest thanks go to my family. I am so lucky to have the four of you in my corner. Jonathan and Andrew, you are the coolest guys I know. Imitating you two has served me well. Mom, you epitomize strength. Thank you for your infinite love. Dad, you have always gone to incredible lengths to support me. Thanks for always making time to talk and for Dad's Track of the Day.

A special thanks goes to those who helped my research directly: to Michel Ledoux for pointing us towards the Bakry-Emery condition, to Mike Noyes for telling me what to expect before it happened, to Florian Sobieczky for feedback, and to Will "GGG" Stanton for so much time and energy down the stretch.

To my advisor, Brian Rider, I cannot thank you enough for working with me. You always knew exactly what I needed. Because of your guidance, I work harder and think smarter than I ever have in my life.

To the Suddath family, whose generosity has no bounds, thank you for welcoming me into your lives. Most importantly, thank you to Kirsten. You, more than anyone, saw me through the struggles I had. You were there when I needed you, and you never wavered. I am so happy that I get to share this with you.

Finally, thanks to all the amazing friends I have made while at CU, you played a fundamental role in my happiness over the past six years. Thank you, and I wish you the best.

Contents

Chapter

Chapter 1

Introduction

Although Random Matrix Theory has its roots in mathematical statistics in the 1920's, the subject did not evolve significantly until the 1950's. This change came because Eugene Wigner, a mathematical physicist, proposed that the eigenvalues of certain random matrices could be used to model the energy levels of highly excited states of heavy nuclei. Even though Wigner's work drew the attention of mathematical physicists, it was not until the 1990's that Random Matrix Theory attracted a wider range of mathematician. The major breakthrough that brought about this change was the discovery of a new class of probability distributions, the Tracy-Widom Laws, named after Craig Tracy and Harold Widom who made the discovery. These laws, which will be discussed in more detail shortly, show up when studying the spectrum of random matrices with specific conditions on the entries. Placing a measure on the space of these matrices makes up what is called a beta ensemble.

In Random Matrix Theory, the most studied beta ensembles are the three Gaussian ensembles, known as the Gaussian Orthogonal Ensemble (GOE), the Gaussian Unitary Ensemble (GUE), and he Gaussian Symplectic Ensemble (GSE). All three are described by a Gaussian measure on the space of $n \times n$ matrices with real Gaussian entries. The GOE corresponds to the case when $\beta = 1$ because the matrices are required to be symmetric with Gaussian entries. The GUE corresponds to the case when $\beta = 2$ due to the fact that the matrices are Hermitian with complex Gaussian entries. Naturally, the GSE corresponds to the case when $\beta = 4$ because of the requirement that the matrices are self-dual matrices with quaternian Gaussian entries. In all three cases, the entries are independent save for the symmetry conditions on the matrices.

The three Gaussian ensembles have well known eigenvalue joint density functions. These are given by

$$\mathbb{P}\left(\lambda_1, \lambda_2, \ldots, \lambda_n\right) = \frac{1}{Z_{\beta,n}} e^{-\frac{\beta}{2}\sum_{k=1}^{n}\lambda_k^2} \prod_{1\leq i<k\leq n} |\lambda_j - \lambda_i|^{\beta} \tag{1.0.1}$$

where $\beta = 1$ for GOE, $\beta = 2$ for GUE, and $\beta = 4$ for GSE. The normalization constant, $Z_{\beta,n}$ can be explicitly computed.

1.1 The Tracy-Widom Laws

The Tracy-Widom distribution was first defined to be the limiting distribution of the properly scaled, largest eigenvalue of the GUE. In other words, if we denote the largest eigenvalue of the GUE as $\lambda_{\max,2}$, then

$$n^{1/6}\left(\lambda_{\max,2} - 2\sqrt{n}\right) \implies TW_2 \tag{1.1.1}$$

where TW_2 is the distribution found by Tracy and Widom in [23]. The idea to center by $2\sqrt{n}$ comes from the following theorem of Wigner, which gives a global picture of the behavior of the GUE eigenvalues.

Theorem 1 (Wigner's Semicircle Law)**.** *Let $\lambda_1^n \leq \lambda_1^n \leq \cdots \leq \lambda_n^n$ denote the ordered eigenvalues of $\frac{1}{\sqrt{n}}X_n$, for X_n a GOE or GUE matrix. Then, for almost every sequence $\{X_n\}_{n=1}^{\infty}$*

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{\lambda_i^n} \implies SC$$

where SC is the probability distribution on $\mathbb{R}$ with density

$$\sigma(x) = \frac{1}{2\pi}\sqrt{4 - x^2}, \quad x \in [-2, 2]$$

This is called the semicircle distribution.

Shortly after their GUE result, Tracy and Widom were able to establish analogous results in the GOE and the GSE cases using the same centering and scaling (see [25]). In other words,

$$n^{1/6}\left(\lambda_{\max,\beta} - 2\sqrt{n}\right) \implies TW_{\beta}. \tag{1.1.2}$$

for $\beta = 1, 2$, and 4.

Sticking to the $\beta = 2$ case for the moment, the Tracy-Widom distribution can be expressed in terms of a Fredholm determinant. More specifically,

$$TW_2(s) = \det(I - A_s)$$

where A_s operates on $L^2(s, \infty)$ with kernel

$$A(x, y) = \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x - y}.$$

Here, Ai is the Airy function defined to be

$$Ai(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{1}{3}t^3 + xt\right) dt.$$

The Tracy-Widom distribution has the integral representation

$$TW_2(s) = \exp\left\{-\int_s^\infty (x - s)q^2(x)dx\right\} \tag{1.1.3}$$

where q is a solution to Painlevé II

$$q''(x) = q^3(x) + xq(x).$$

Interestingly enough, q is asymptotically similar to the Airy function, or in other words, there are constants $c_1, c_2 > 0$ such that

$$c_1 Ai(x) \leq q(x) \leq c_2 Ai(x).$$

In the $\beta = 1$ and $\beta = 4$ cases, the results of [25] include integral representations similar to 1.1.3, the one for $\beta = 2$. Solutions to Painlevé II can be numerically approximated, which has allowed TW_β to be tabulated for $\beta = 1, 2$, and 4.

In 1999, the Tracy-Widom distribution appeared in the work of Baik, Deift, and Johansson while studying the length of the longest increasing subsequence of a random permutation of n numbers. More specifically, in [3], the authors proved that as $n \to \infty$

$$n^{-1/6}\left(l_n - 2\sqrt{n}\right) \implies TW_2, \tag{1.1.4}$$

where l_n is the length of the longest increasing subsequence of a permutation chosen uniformly from S_n, the symmetric group over n elements. The scaling of $n^{-1/6}$ tells us that l_n behaves just like $n^{1/3}\lambda_{\max,2}$ as $n \to \infty$.

Classically, the β ensembles were studied for only $\beta = 1, 2,$ and 4. The classical results are commonly due to the fact that the Gaussian ensembles are amenable to direct computation. The finite dimensional correlation functions can all be expressed explicitly in terms of Hermite polynomials. In turn, the asymptotics of these polynomials can be studied with combinatorics, complex analysis, integrable systems and probability theory. Still, 1.0.1 has physical applications for all $\beta > 0$. This ensemble can be used to model a one-dimensional Coulomb gas with inverse temperature β, and it is tied to Calogero-Sutherland quantum systems.

A major breakthrough in the study of these β ensembles was provided by Dumitriu and Edelman in [5]. They discovered the existence of a fairly simple tridiagonal matrix model for all β, the eigenvalues of which have 1.0.1 as their joint probability distribution. This matrix, called the Hermite tridiagonal matrix, is given by

$$H_{\beta,n} = \begin{bmatrix} g_1 & \chi_{\beta(n-1)} & & & & & \\ \chi_{\beta(n-1)} & g_2 & \chi_{\beta(n-2)} & & & & \\ & \chi_{\beta(n-2)} & g_3 & \chi_{\beta(n-3)} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & \chi_{\beta 2} & g_{n-1} & \chi_\beta \\ & & & & & \chi_\beta & g_n \end{bmatrix},$$

where $g_1, g_2, \ldots, g_n$ are independent with $g_i \sim N(0,2)$, and $\chi_\beta, \chi_{\beta 2}, \ldots, \chi_{\beta(n-2)}$ are independent χ random variables.

This matrix model has already had major implications in the study of the distributional limits of the eigenvalues in the β-Hermite ensemble. Heuristically, Sutton [22] and Edelman and Sutton [7] argued that the rescaled tridiagonal matrix model could be associated with a continuum operator. More specifically, they provided hope that limiting distributions for general β could be found by associating the rescaled Hermite tridiagonal matrix

$$n^{1/6}\left(H_{\beta,n} - 2\sqrt{n} I_n\right)$$

with the stochastic Airy operator

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

Here b' is the formal derivative of a standard Brownian motion, white noise. In 2006, Ramírez, Rider, and Virág rigorously proved this conjecture, and in the process, extended the definition to the Tracy-Widom Laws to $\beta > 0$. The following definition is consistent with all prior definitions of TW_β.

Definition 2 (Tracy-Widom Distribution)**.** Let $x \to b(x)$ be a standard Brownian motion. For $\beta > 0$ define the general Tracy-Widom law to be

$$TW_\beta = \sup_{f \in L} \left\{ \frac{2}{\sqrt{\beta}} \int_0^\infty f^2(x) db(x) - \int_0^\infty \left[(f'(x))^2 + x f^2(x) \right] dx \right\}, \tag{1.1.5}$$

where L is the space of functions which vanish at the origin and satisfy

(1) $\int_0^\infty f^2(x)dx = 1$,

(2) $\int_0^\infty \left[(f'(x))^2 + x f^2(x)\right] dx < \infty$.

Much of the work surrounding the Tracy Widom distributions takes place in the continuum, but it is natural to ask questions regarding the eigenvalues before the limit is taken. Understanding the rate of convergence of the largest eigenvalue to TW_β is important when considering the use of TW_β to make approximations in finite dimensional models. This rate of convergence is often referred to as "Small Deviations", and it is the focus of this thesis.

1.2 Organization

This dissertation establishes small deviation inequalities for the largest eigenvalue of the β-Jacobi ensemble 4.0.1. The main results are Theorem 13 and Theorem 14, and Corollary 15 follows immediately. Chapter 2 discusses some of the major results that aided in connecting finite random matrix theory to infinite random matrix theory . These include the tridiagonal matrix models of Dimitriu and Edelman and the work of Ramírez, Rider, and Virág to extend the Tracy Widom Laws to $\beta > 0$. Chapter 3 gives a more detailed description of the topic of small deviations including a brief overview of previous work.

Chapter 4 is intended to provide the reader with preliminary information regarding the β-Jacobi ensemble, and it is also where the results of this dissertation are stated. In Chapter 5, the connection between the variational picture with the finite dimensional Jacobi matrix model is made clear by proving Lemma 16, which plays a major role in the proof of Theorems 13 and 14. Chapter 6 is devoted to the proof of Theorem 13, and Chapter 7 focuses on the proof of Theorem 14. An immediate consequence of these theorems is Corollary 15, which is proved in Chapter 8.

Chapter 2

Background

2.1 The Laguerre Ensemble

The so called beta ensembles are point processes on $\mathbb{R}$ that are defined, for $\beta > 0$, by n-level joint density functions. The three most common are the Hermite, Laguerre, and Jacobi ensembles. The Hermite ensemble, defined by 1.0.1, was discussed in the introduction, and the Jacobi ensemble will be discussed in Chapter 4. The Laguerre ensemble is defined by

$$P(\lambda_1, \lambda_2, \ldots, \lambda_n) = \frac{1}{Z_{\beta,n}} e^{-\frac{\beta}{2}\sum_{k=1}^{n}\lambda_k} \prod_{k=1}^{n} \lambda_k^{\frac{\beta}{2}(a+1)-1} \prod_{1\le j<k\le n} |\lambda_j - \lambda_k|^{\beta}. \tag{2.1.1}$$

This joint density is the joint eigenvalue density of the collection of $n \times n$ Wishart matrices. These are matrices of the form $W_n = A_n A_n^T$, where A_n is an $n \times M(n)$ matrix (often $M(n) = n + a$) with i.i.d. entries of mean zero, variance $1/n$ satisfying certain moment conditions.

In the Laguerre ensemble, the result analogous to Wigner's Semicircle Law is the Marchenko-Pastur Distribution.

Theorem 3 (Marchenko-Pastur Distribution)**.** *Let $0 < \lambda_1^n < \lambda_2^n < \cdots < \lambda_n^n$ denote the ordered eigenvalues of a Wishart matrix, W_n. Then, with the convergence being weakly, in probability*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{N} \delta_{\lambda_i^n} = F_a,$$

where F_a is a distribution function with density

$$f_a(x) = \frac{\sqrt{(x-b_-)(b_+-x)}}{2\pi x} \mathbb{1}_{[b_-,b_+]}(x)$$

with $b_- = (1-\sqrt{a})^2$ and $b_+ = (1+\sqrt{a})^2$.

2.2 Tridiagonal Matrix Models

In 2002, Ioana Dumitriu and Alan Edelman introduced two fairly simple matrix models whose joint eigenvalue densities corresponded to the Hermite and Laguerre ensembles for general β. These matrix models opened the door to new approaches to studying the spectrum of random matrices. In 2005, Brian Sutton produce an analagous matrix model for the Jacobi ensemble. These matrix models will be discussed here, but first it may be useful to recall the definitions of normal, chi, and beta random variables.

Definition 4. A random variable is said to be *normally distributed* with mean μ and variance σ^2 if it has the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}.$$

A random variable is said to be χ distributed with k degrees of freedom if it has the probability density function

$$f(x) = \frac{2^{1-\frac{k}{2}} x^{k-1} e^{\frac{x^2}{2}}}{\Gamma(-\frac{k}{2})}, \quad x \in [0, \infty)$$

where $\Gamma(z)$ is the gamma function. A random variable is said to be *beta distributed* with parameters s and t if it has the probability density function

$$\frac{\Gamma(s+t)}{\Gamma(s)\Gamma(t)} x^{s-1}(1-x)^{t-1}, \quad x \in [0, 1].$$

Let $g_1, g_2, \ldots, g_n$ be $N(0, 2)$ random variables, and let $\chi_\beta, \chi_{\beta 2}, \ldots, \chi_{\beta(n-1)}$ be χ random variables with the given degrees of freedom. For $\beta > 0$, define the $n \times n$ Hermite tridiagonal matrix as

$$H_{n,\beta} = \frac{1}{\sqrt{\beta}} \begin{bmatrix} g_1 & \chi_{\beta(n-1)} & & & & & \\ \chi_{\beta(n-1)} & g_2 & \chi_{\beta(n-2)} & & & & \\ & \chi_{\beta(n-2)} & g_3 & \chi_{\beta(n-3)} & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & \chi_{\beta 2} & g_{n-1} & \chi_\beta \\ & & & & & \chi_\beta & g_n \end{bmatrix}. \quad (2.2.1)$$

The matrix $H_{n,\beta}$ is symmetric and all random variables are independent except for the dependence imposed by the symmetry condition. In [5], Dimitriu and Edelman established the following theorem.

Theorem 5 (Edelman-Dimitriu)**.** *For any $\beta > 0$, the Hermite tridiagonal matrix has joint eigenvalue density given by 1.0.1.*

Also in [5], Dimitriu and Edelman establish a similar result for the Laguerre ensemble. Let $\chi_{\beta(a+1)}, \chi_{\beta(a+2)}, \ldots, \chi_{\beta(a+n)}$ and $\widetilde{\chi}_{\beta}, \widetilde{\chi}_{\beta 2}, \ldots, \widetilde{\chi}_{\beta(n-1)}$ be χ random variables with the given degrees of freedom. For any $\beta > 0$ and $a > -1$, define the $n \times n$ Laguerre tridiagonal matrix as $L_{n,a,\beta} = A_{n,a,\beta} \cdot A^T_{n,a,\beta}$ where

$$A_{n,a,\beta} \sim \begin{bmatrix} \chi_{\beta(a+n)} & & & & \\ \widetilde{\chi}_{\beta(n-1)} & \chi_{\beta(a+n-1)} & & & \\ & \widetilde{\chi}_{\beta(n-1)} & \chi_{\beta(a+n-2)} & & \\ & & \ddots & \ddots & \\ & & & \widetilde{\chi}_{\beta} & \chi_{\beta(a+1)} \end{bmatrix}.$$

The matrix $L_{n,a,\beta}$ is symmetric and all χ random variables are independent. Unlike in the Hermite case, the entries of $L_{n,a,\beta}$ are not independent.

Theorem 6 (Edelman-Dimitriu)**.** *For any $\beta > 0$ and $a > -1$, the Laguerre tridiagonal matrix has joint eigenvalue density given by 2.1.1.*

2.3 Tracy-Widom for general β

Classical results regarding the limiting eigenvalue distributions relied heavily on the fact that $\beta = 1, 2, 4$ produces integrable systems. The tridiagonal matrix models provided by Edelman and Dimitriu were a major breakthrough in the effort to generalize the Tracy-Widom laws to all $\beta > 0$. In [7], Edelman and Sutton used these matrix models to make the following two conjectures.

Conjecture 1 (Edelman-Sutton). *As $n \to \infty$, the centered and scaled Hermite tridiagonal matrix*

$$H_n = n^{1/6}(H_{n,\beta} - 2\sqrt{n})$$

converges to the Stochastic Airy Operator

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

Here b'_x is "white noise", the formal derivative of Brownian motion.

Conjecture 2 (Edelman-Sutton). *Let σ_k denote the k-th smallest singular value of the bidiagonal matrix $B_{\beta,a}$. As $n \to \infty$, the family of rescaled singular values $\{\sqrt{n}\sigma_k\}$ converges in law to the singular values of the following random differential operator*

$$\mathcal{L}_{\beta,a} = -\sqrt{x}\frac{d}{dx} + \frac{a}{2\sqrt{x}} + \frac{1}{\sqrt{\beta}} b'(x). \tag{2.3.1}$$

These conjectures are based on two similar heuristic arguments that rely heavily on the tridiagonal matrix models provided in [5]. Due to the similarity of the arguments, only Conjecture 1 will be discussed. The basic idea behind Conjecture 1 was to center by $2\sqrt{n}$ because of Theorem 1, and then look for an appropriate scaling factor. In other words, one hopes that there is some γ such that the quantity

$$\widehat{H}_n := n^\gamma \left(H_{n,\beta} - 2\sqrt{n} I_n\right) \tag{2.3.2}$$

converges to the stochastic Airy operator $\mathcal{H}_\beta$. By the Central Limit Theorem, a χ random variable with t degrees of freedom is asymptotically like $\sqrt{t} + G/\sqrt{2}$ where $G \sim N(0,1)$. Using the Hermite tridiagonal and the asymptotics just mentioned, 2.3.2 looks like the discrete analog of the Stochastic Airy Operator. Since $\mathcal{H}_\beta$ operates on suitably nice functions ϕ, so it makes sense to think of $\widehat{H}_n$ as operating on a discretized version of ϕ. Towards this end, write $k = \lfloor xn^\alpha \rfloor$ for some α, and denote $\phi_k = \phi(x)$. Taylor expanding of ϕ leads to the appropriate choices for α and γ. For more details, see Chapter 2 of [18].

Considering the fact that Brownian motion is nowhere differentiable, proving these conjectures is not straightforward. Nonetheless, both conjectures were proved in 2006 by Ramírez, Rider, and Virág. The statement of their result is the following.

Theorem 7 (Ramírez, Rider, and Virág)**.** *With probability one, for each $k \geq 0$ the set of eigenvalues of $\mathcal{H}_\beta$ has a well defined $(k+1)$st lowest element Λ_k. Moreover, let $\lambda_{\beta,1} \geq \lambda_{\beta,2} \geq \cdots$ denote the eigenvalues of the Hermite β-ensemble H_n^β. Then the vector*

$$\left(n^{1/6}\left(2\sqrt{n} - \lambda_{\beta,l}\right)\right)_{l=1,\ldots,k}$$

converges to $(\Lambda_0, \Lambda_1, \ldots, \Lambda_{k-1})$ in distribution, as $n \to \infty$.

Chapter 3

Small Deviations

Recall that the Tracy-Widom Theorem for the largest eigenvalue of the GUE reads

$$n^{1/6}(\lambda_{\max}(H_{2,n}) - 2\sqrt{n}) \implies TW_2.$$

To motivate the topic of small deviations, it is more useful to rewrite the limit theorem as

$$\lim_{n\to\infty} \mathbb{P}\left(\lambda_{\max}(H_{2,n}) \leq 2\sqrt{n}(1 + sn^{-2/3})\right) = F_{TW_\beta}(s). \tag{3.0.1}$$

Given that the known shape of the TW_β is

$$\mathbb{P}\left(TW_\beta \leq t\right) \sim e^{\beta t^3/24} \text{ as } t \to -\infty, \text{ and } \mathbb{P}\left(TW_\beta \geq t\right) \sim e^{-2\beta t^{3/2}/3} \text{ as } t \to \infty,$$

one would hope that, for small ε, there exits a constant $C > 0$ such that

$$\mathbb{P}\left(\lambda_{\max} \leq 2\sqrt{n}(1-\varepsilon)\right) \leq Ce^{-\beta\varepsilon^3 n^2/C}, \tag{3.0.2}$$

$$\mathbb{P}\left(\lambda_{\max} \geq 2\sqrt{n}(1+\varepsilon)\right) \leq Ce^{-\beta\varepsilon^{3/2} n/C}, \tag{3.0.3}$$

for all $n \geq 1$. The need for ε to be small is to encompass the $sn^{-2/3} \to 0$ regime, and it is precisely why this topic is called "small deviations".

The small deviation upper bounds 3.0.2 and 3.0.3 are not the only small deviation inequalities. One would hope that the upper bounds are tight, or in other words, that there is a different constant $C > 0$ such that

$$\mathbb{P}\left(\lambda_{\max} \leq 2\sqrt{n}(1-\varepsilon)\right) \geq Ce^{-\beta\varepsilon^3 n^2/C},$$

$$\mathbb{P}\left(\lambda_{\max} \geq 2\sqrt{n}(1+\varepsilon)\right) \geq Ce^{\beta\varepsilon^{3/2} n/C},$$

for all $n \geq 1$ and $\varepsilon > 0$.

In contrast, the topic of large deviations concerns itself with ε larger than $O(1)$. Consider the right-tail limit from [2]

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}\left(\lambda_{\max}(H_{2,n}) \geq 2\sqrt{n}(1+\varepsilon)\right) = -J_{GUE}(\varepsilon) \tag{3.0.4}$$

where, for every $\varepsilon > 0$,

$$J_{GUE}(\varepsilon) = 4\int_0^{\varepsilon} \sqrt{x(x+2)}dx.$$

When ε is small, $J_{GUE}(\varepsilon)$ is of order $\varepsilon^{3/2}$, whereas, for larger ε, $J_{GUE}(\varepsilon)$ is of order ε^2. Hence, one would expect a large deviation right-tail inequality of the form

$$\mathbb{P}\left(\lambda_{\max} \geq 2\sqrt{n}(1+\varepsilon)\right) \leq Ce^{-\beta\varepsilon^2 n/C}.$$

for all $n \geq 1$. This inequality follows from standard net arguments on the corresponding Gaussian matrices (see e.g. [14])

3.1 Classical small deviation results

Classical results ($\beta = 1, 2, 4$) in the area of small deviations was rather fragmentary until Ledoux and Rider establish small deviation results for general β in [16]. Perhaps the most extensive work for classical ensembles was done by Michel Ledoux in [13], where a right tail upper bound was established for the GUE, LUE, and JUE cases. In [15], Ledoux also proves a recurrence relation for the GOE leading to a right tail small deviation inequality. The following theorem is the statement for the GUE.

Theorem 8 (Proposition 5.2 from [13])**.** *For every* $0 < \varepsilon \leq 1$ *and* $n \geq 1$

$$\mathbb{P}\left(\lambda_{\max}(H_{2,n}) \geq 2\sqrt{n}(1+\varepsilon)\right) \leq Ce^{-n\varepsilon^{3/2}/C}$$

where $C > 0$ *is a numerical constant.*

The proof of this result (as well as in the LUE and JUE cases) requires a recurrence formula, also established in [13], for the moments of the mean spectral measures ($\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n \delta_{\lambda_i^n}\right)$). In the

GUE case, these moments are given by

$$b_p = b_p^n = \int (\sigma x)^{2p} \frac{1}{n} \sum_{l=0}^{n-1} P_l^2 d\mu, \qquad\qquad p \in \mathbb{n},$$

where $b_0 = 1, b_1 = \sigma^2 n^2$, and the P_l are the Hermite orthogonal polynomials. The recursion formula then reads

$$b_p = 4n\sigma^2 \frac{2p-1}{2p+2} b_{p-1} + 4\sigma^4 p(p-1) \frac{2p-1}{2p+2} \cdot \frac{2p-3}{2p} b_{p-2} \tag{3.1.1}$$

for every $p \geq 1$. A simple induction argument on 3.1.1 gives an upper bound for b_p involving the moments of Wigner's semicircle law. The asymptotics of these moments yields Theorem 8.

It should also noted that in [13, 14], Ledoux discusses the fact that Theorem 8 can be shown to follow from the results of Johannson [11] for a more general model. For a more complete discussion on classical small deviation inequalities see [14].

3.2 Small deviation results for general β

In [16], Brian Rider and Michel Ledoux made considerable strides in the effort to prove small deviation inequalities for general β-ensembles. In the β-Hermite case, they established tight bounds for both the left and right tails. The left tail and right tail upper bounds are as follows.

Theorem 9 (Theorem 1 from [16])**.** *For all $n \geq 1$, $0 < \varepsilon \leq 1$ and $\beta \geq 1$:*

$$\mathbb{P}\left(\lambda_{\max}(H_{\beta,n} \geq 2\sqrt{n}(1+\varepsilon))\right) \leq Ce^{-\beta n \varepsilon^{3/2}/C},$$

and

$$\mathbb{P}\left(\lambda_{\max}(H_{\beta,n} \geq 2\sqrt{n}(1-\varepsilon))\right) \leq C^{\beta} e^{-\beta n^2 \varepsilon^3/C},$$

where C is a numerical constant.

The following theorem established tightness by stating the optimal lower bounds for both tails.

Theorem 10 (Theorem 4 from [16])**.** *There is a numerical constant C so that*

$$\mathbb{P}\left(\lambda_{\max}(H_{\beta,n}) \geq 2\sqrt{n}(1+\varepsilon)\right) \geq C^{-\beta} e^{-C\beta n \varepsilon^{3/2}},$$

and

$$\mathbb{P}\left(\lambda_{\max}(H_{\beta,n}) \geq 2\sqrt{n}(1-\varepsilon)\right) \geq C^{-\beta}e^{-C\beta n^2\varepsilon^3}.$$

The first inequality holds for all $n > 1$, $0 < \varepsilon \leq 1$ and $\beta \geq 1$. For the second inequality, the range must be kept sufficiently small, $0 < \varepsilon \leq 1/C$.

Also in [16], Ledoux and Rider established small deviation upper bounds for the Laguerre ensemble as well as a lower bound for the right-tail. In the Hermite case, the left-tail lower bound required a Gaussian argument, which was not available in the Laguerre case.

3.3 Motivation from the continuum

The Tracy-Widom law, established in [20], is identified via a random variation principle. In particular, it holds that

$$TW_\beta = \sup_{f\in\mathcal{L}^*}\left\{\frac{2}{\sqrt{\beta}}\int_0^\infty f^2(x)db(x) - \int_0^\infty \left[(f'(x))^2 + xf^2(x)\right]dx\right\}, \tag{3.3.1}$$

where $x \mapsto b(x)$ is a standard Brownian motion, and $\mathcal{L}^*$ is defined below. The small deviation results of Ledoux and Rider in [16] are achieved by retooling (for finite n) the techniques used to prove the general β Tracy-Widom theorem. For that reason, we will discuss some of the details from their proof in this section and how they shed light on the topic of small deviations.

As mentioned in the previous chapter (see 1), the Edelman-Sutton Conjecture suggested that the appropriately scaled Hermite tridiagonal matrix can be viewed as the finite dimensional analog of the stochastic Airy operator

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

One would then hope that limiting eigenvalue distributions can be obtained by studying the stochastic Airy operator.

At first glance, this operator immediately poses a problem. Since Brownian motion is nowhere differentiable, white noise is only defined formally. Towards establishing a proper framework for studying $\mathcal{H}_\beta$ let us view b'_x as a generalized function. Let ϕ be a smooth function of compact support on $(0,\infty)$. Integrating by parts yields

$$\int_0^y b'_x \phi(x)dx = b_y\phi(y) - \int_0^y b_x\phi'(x)dx,$$

for which the right hand side is a continuous function in y having $b'_y\phi(y)$ as its derivative.

We now turn to the eigenvalue problem, starting first with the eigenvalues of $\mathcal{H}_\beta$. An eigenvalue and eigenfunction pair of $\mathcal{H}_\beta$ is an ordered pair, $(\lambda, f) \in \mathbb{R} \times \mathcal{L}_*$, that satisfies $\mathcal{H}_\beta f = \lambda f$ in the sense of distributions. This can be rewritten as

$$f''(x) = (x - \lambda + \frac{2}{\sqrt{\beta}}b'_x)f.$$

After integrating by parts, the above equality reads

$$\int_0^\infty \phi''(x)f(x)dx = \int_0^\infty (x-\lambda)\phi(x)f(x)dx + \int_0^\infty \frac{2}{\sqrt{\beta}}\left[\int_0^x b_y f'(y)dy - b_x f(x)\right]\phi'(x)dx$$

when viewed in the distributional sense. This weak notion of an eigenvalue starts to bring the variational characterization of the eigenvalue problem into view.

If ϕ is smooth in the sense of Schwarz distributions, then $-\mathcal{H}_\beta\phi$ can be applied to ϕ as a linear functional to get the following quadratic form

$$\prec \phi, \mathcal{H}_\beta\phi \succ := \int_0^\infty (\phi'(x))^2 dx + \int_0^\infty x\phi^2(x) - \frac{2}{\sqrt{\beta}}\int_0^\infty b'_x\phi^2(x)dx \tag{3.3.2}$$

on C_0^∞. This quadratic form can be extended to operate on the Hilbert space

$$\mathcal{L}^* := \left\{ f : f(0) = 0, \text{ and } \int_0^\infty (f')^2 + (1+x)^2 f^2 dx \right\}$$

equipped with the norm $\|f\|_*^2 = \int_0^\infty (f')^2 + (1+x)^2 f^2 dx$. This extension helps to avoid technical difficulties, and $\mathcal{L}^*$ can even be weakened slightly.

The smallest eigenvalue of $\mathcal{H}_\beta$ can then be characterized by the variational principle

$$\widetilde{\Lambda}_0 := \inf_{f\in\mathcal{L}_*} \left\{ \prec f, \mathcal{H}_\beta f \succ : f(0) = 0 \text{ and } \|f\|_2 = 1 \right\}.$$

Notice that this is the same as 3.3.1. As one might expect, $\Lambda_0 = \widetilde{\Lambda}_0$ where Λ_0 is the smallest eigenvalue of $\mathcal{H}_\beta$.

The next step is to connect the eigenvalues of H_n to the eigenvalues of $\mathcal{H}_\beta$. A natural approach is to see if the eigenvalues of H_n could be characterized by a "discrete" variational principle. For $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$, define the quadratic form

$$n^{-1/6}\langle v, v\rangle_{H_n} := v^T H_n v.$$

Recalling the definition of the Hermite tridiagonal 2.2.1, we get

$$\begin{aligned}\langle v, v\rangle_{H_n} &= -n^{1/3}\sum_{i=1}^{n}(v_k - v_{k+1})^2 - \frac{2}{n^{1/6}}\sum_{i=1}^{n}(\sqrt{n} - \frac{1}{\sqrt{\beta}}\mathbb{E}\chi_{\beta(n-k)})v_k v_{k+1} \\ &\quad + \frac{1}{n^{1/6}\sqrt{\beta}}\sum_{i=1}^{n} g_k v_k^2 + \frac{2}{n^{1/6}}\sum_{i=1}^{n}(\frac{1}{\sqrt{\beta}}\chi_{\beta(n-k)} - \mathbb{E}\chi_{\beta(n-k)})v_k v_{k+1}.\end{aligned}$$

The idea here is to that these sums look like discretized version of the integrals in 3.3.2. To make this connection formal, H_n needs to operator on $\mathcal{L}_*$. This can be done by identifying $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ with a function in $L^2(\mathbb{R}^+)$ by defining the step function $v(x) = v_k$ for $x \in [(k-1)n^{1/3}, kn^{1/3})$ and $v(x) = 0$ for $x > n^{2/3}$.

From here, the idea is to use the discrete quadratic form to show that the largest eigenvalue of H_n converges to the Λ_0. Because of this, the quadratic form can be used to find sharp, shape estimates on the largest eigenvalue for finite n, which is the goal when trying to find small deviation inequalities.

Chapter 4

The Jacobi Ensemble

The Jacobi beta ensemble is the point process on $\mathbb{R}$ defined by the n-level joint density: for $\beta > 0$

$$P(\lambda_1, \lambda_2, \ldots, \lambda_n) = \frac{1}{Z_{\beta,n}} \prod_{k=1}^{n} \lambda_k^{\frac{\beta}{2}(a+1)-1} (1-\lambda_k)^{\frac{\beta}{2}(b+1)-1} \prod_{1 \le j < k \le n} |\lambda_j - \lambda_k|^{\beta}. \tag{4.0.1}$$

As in the Hermite and Laguerre cases, a tridiagonal matrix model exists and is an invaluable tool in studying the Jacobi ensemble for general β.

4.1 The Jacobi Tridiagonal Matrix

The Jacobi tridiagonal matrix is given by $J_{\beta,n,a,b} = B_{\beta,n,a,b} \cdot B_{\beta,n,a,b}^T$, where

$$B_{\beta,n,a,b} = \begin{bmatrix} c_n & -s_n c'_{n-1} & & & \\ & c_{n-1}s'_{n-1} & -s_{n-1}c'_{n-2} & & \\ & & c_{n-2}s'_{n-2} & \ddots & \\ & & & \ddots & -s_2 c'_1 \\ & & & & c_1 s'_1 \end{bmatrix}. \tag{4.1.1}$$

where the c_i's and c'_i's are independent random variables defined by

$$c_i \sim \sqrt{\text{Beta}\left(\tfrac{\beta}{2}(an+i), \tfrac{\beta}{2}(bn+i)\right)}, \quad c'_i \sim \sqrt{\text{Beta}\left(\tfrac{\beta}{2}i, \tfrac{\beta}{2}(an+bn+1+i)\right)}$$

and the s_i's and s'_i's are defined by

$$s_i = \sqrt{1-c_i^2}, \quad s'_i = \sqrt{1-c_i'^2}.$$

The following theorem, established in [22], encompasses the importance of the Jacobi tridiagonal matrix

Theorem 11 (Sutton)**.** *For any $\beta > 0$ and $a, b > -1$, the Jacobi tridiagonal matrix has joint eigenvalue density given by 4.0.1.*

The limiting distribution of the largest eigenvalue of the JUE was first discovered by Collins in [4], where the largest eigenvalue was shown to converge in distribution to

$$\begin{aligned} \gamma &= \frac{(a+1)(a+b+1)}{(a+b+2)^2} + \frac{(b+1)}{(a+b+2)^2} + \frac{2\sqrt{(a+1)(a+b+1)(b+1)}}{(a+b+2)^2} \\ &= \left(\frac{\sqrt{(a+1)(a+b+1)}}{a+b+2} + \frac{\sqrt{b+1}}{a+b+2} \right)^2 \end{aligned} \tag{4.1.2}$$

Remark 12. In [4], the work is actually conducted for the Jacobi Ensemble with support on $[-1, 1]$. In this paper, the transformation $x \mapsto \frac{x+1}{2}$ has been used for convenience.

4.2 Properties of the Beta Random Variable

A random variable X is said to be beta distributed with shape parameters $s, t > 0$, written $X \sim \text{Beta}(s, t)$, if it has the probability density function

$$\begin{aligned} f(x) &= \frac{x^{s-1}(1-x)^{t-1}}{\int_0^1 x^{s-1}(1-x)^{t-1}dx} \\ &= \frac{\Gamma(s+t)}{\Gamma(s)\Gamma(t)} x^{s-1}(1-x)^{t-1} \end{aligned}$$

where $x \in (0, 1)$ and $\Gamma(z)$ is the well known gamma function defined by

$$\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} dx.$$

The mean of X is then given by

$$\mathbb{E}\, X = \frac{s}{s+t}, \tag{4.2.1}$$

and due to the symmetry $1 - X \sim \text{Beta}(t, s)$.

To see why the gamma function shows up write

$$\begin{aligned}\Gamma(s)\Gamma(t) &= \int_0^\infty e^{-x}x^{s-1}dx \int_0^\infty e^{-y}y^{t-1}dy \\ &= \int_0^\infty \int_0^\infty e^{-x-y}x^{s-1}y^{t-1}dxdy.\end{aligned}$$

Changing variables to $x = zr, y = z(1-r)$ yields

$$\begin{aligned}\Gamma(s)\Gamma(t) &= \int_0^\infty \int_0^1 e^{-z}z^{s+t-2}r^{s-1}(1-r)^{t-1}z\,drdz \\ &= \int_0^\infty e^{-z}z^{s+t-1}dz \int_0^1 r^{s-1}(1-r)^{t-1}dr \\ &= \Gamma(s+t)\cdot \int_0^1 r^{s-1}(1-r)^{t-1}dr\end{aligned}$$

as desired.

Due to the nature of the random variables in 4.1.1, it will be important to have bounds on $\mathbb{E}\sqrt{X}$ and $\mathbb{E}\sqrt{X(1-x)}$. Towards this end we have

$$\begin{aligned}\mathbb{E}\sqrt{X} &= \frac{\Gamma(s+t)}{\Gamma(s)\Gamma(t)} \int_0^1 \sqrt{x}x^{s-1}(1-x)^{t-1}dx \\ &= \frac{\Gamma(s+t)}{\Gamma(s+t-1/2)} \cdot \frac{\Gamma(s-1/2)}{\Gamma(s)} \cdot \mathbb{E}Y\end{aligned}$$

where $Y \sim \text{Beta}(s-1/2, t)$. Now, recall that a χ_k random variable has mean

$$\mathbb{E}\chi_k = \sqrt{2}\cdot \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{k}{2})},$$

and thus

$$\mathbb{E}\sqrt{X} = \frac{\mathbb{E}\chi_{2(s+t)-1}}{\mathbb{E}\chi_{2s-1}} \cdot \frac{2s-1}{2(s+t)-1}. \tag{4.2.2}$$

A similar calculation yields

$$\mathbb{E}\sqrt{X(1-X)} = \frac{\mathbb{E}\chi_{2t}}{\mathbb{E}\chi_{2s-1}} \cdot \frac{2s-1}{2(s+t)}. \tag{4.2.3}$$

We can get upper and lower bounds by using the following bounds:

$$\begin{aligned}\mathbb{E}\chi_r &\le \sqrt{r}, & r > 0 \\ \mathbb{E}\chi_r &\ge \sqrt{r-\tfrac{1}{2}}, & r > 1 \\ \mathbb{E}\chi_r &\ge \tfrac{r}{\sqrt{r+1}}, & r > 0.\end{aligned}$$

The first two follow applying Jensen's inequality to $\mathbb{E}\chi_r = \sqrt{2}\frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})}$, and the second follows from

$$x^{1-s} \leq \frac{\Gamma(x+1)}{\Gamma(x+s)} \leq (x+s)^{1-s}$$

which was established in [19]. Using these bounds for positive s and t, one has

$$\frac{\sqrt{2s-1}}{\sqrt{2(s+t)}} \leq \quad \mathbb{E}\sqrt{X} \quad \leq \frac{\sqrt{2s}}{\sqrt{2(s+t)-1}}, \tag{4.2.4}$$

$$\frac{\sqrt{2s-1}}{2(s+t)} \cdot \sqrt{2t-\frac{1}{2}} \leq \mathbb{E}\sqrt{X(1-X)} \leq \frac{\sqrt{2t}\sqrt{2s}}{\sqrt{2(s+t)-1}}, \qquad t > \frac{1}{2} \tag{4.2.5}$$

and

$$\frac{\sqrt{2s-1}}{2(s+t)} \cdot \frac{2t}{\sqrt{2t+1}} \leq \mathbb{E}\sqrt{X(1-X)} \leq \frac{\sqrt{2t}\sqrt{2s}}{\sqrt{2(s+t)-1}}. \tag{4.2.6}$$

4.3 Results

The original contributions of this thesis are the following two theorems and corollary. The first theorem is the right-tail upper bound for the β-Jacobi ensemble, and the second theorem is the left-tail upper bound.

Theorem 13. *Let $a \in [0,\infty)$ and $b \in (0,\infty)$. Then for all $n \geq 1$ and $0 < \varepsilon \leq 1$:*

$$\mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \geq \gamma\sqrt{n}(1+\varepsilon)\right) \leq C_\beta e^{-\beta(a+b)n\varepsilon^{3/2}/C_\beta}, \tag{4.3.1}$$

where C_β is a numerical constant.

Chapter 6 contains the proof for the right-tail upper bound.

Theorem 14. *Let $a \in [0,\infty)$ and $b \in (0,\infty)$. Then for all $n \geq 1$ and $0 < \varepsilon \leq 1$:*

$$\mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \leq \gamma\sqrt{n}(1-\varepsilon)\right) \leq C_\beta e^{-\beta(a+b)\varepsilon^3 n^2/C_\beta} \tag{4.3.2}$$

where C_β is a numerical constant.

The proof of the left-tail upper bound is contained in Chapter 7. In the β-Hermite case, as proved in [16], the Gaussian random variables in the Hermite tridiagonal matrix play an important role in the left-tail upper bound. This argument is unavailable in the β-Jacobi case.

The upper bounds from Theorems 13 and Theorem 14 are enough to produce the following corollary regarding the variance, for finite n, of the largest eigenvalue. This bound is of the expected order, and a short proof is contained in Chapter 8.

Corollary 15. *For $\beta \geq 1$ and $n \in \mathbb{N}$*

$$Var[\lambda_{\max}(J_{\beta,n})] \leq C_\beta n^{-1/3},$$

where the constant C_β depends on β.

Chapter 5

Operator Bound

This chapter begins the proof of Theorems 13 and 14. With the framework from the previous chapter in mind, define

$$\frac{1}{\sqrt{n}} J_n(v) = v^T \left[J_{\beta,n,a,b} - \gamma I_n \right] v.$$

where γ is the appropriate centering from 4.1.2. Dividing by $\sqrt{n}$ makes for better comparison with [16]. From this point forward, the dependence of $J_n(v)$ on n will be suppressed by simply writing $J(v)$. Then

$$\begin{aligned} J(v) \quad = \quad & \sqrt{n} \sum_{k=1}^{n} c_{n-k+1}^2 s_{n-k+1}'^2 \, v_k^2 + \sqrt{n} \sum_{k=1}^{n} c_{n-k}'^2 s_{n-k+1}^2 \, v_k^2 \\ & + 2\sqrt{n} \sum_{k=1}^{n-1} c_{n-k} s_{n-k+1} c_{n-k}' s_{n-k}' \, v_k v_{k+1} - \gamma \sqrt{n} \sum_{k=1}^{n} v_k^2. \end{aligned} \tag{5.0.1}$$

Proving Theorems 13 and 14 now becomes a task of estimating

$$\mathbb{P}\left(\sup\nolimits_{\|v\|_2 = 1} J(v) \geq \varepsilon \gamma \sqrt{n} \right) \quad \text{and} \quad \mathbb{P}\left(\sup\nolimits_{\|v\|_2 = 1} J(v) \leq -\varepsilon \gamma \sqrt{n} \right).$$

The following lemma will play a fundamental role in the rest of the proof, and it will be carried out in the next two subsections.

Lemma 16. *For $\beta \geq 1$, $c > 0$ set*

$$J_c(v) \quad = \quad \sum_{k=1}^{n} Z_k v_k^2 + \sum_{k=1}^{n} \widetilde{Z}_k v_k^2 + 2 \sum_{k=1}^{n-1} Y_k v_k v_{k+1} \tag{5.0.2}$$

$$-c\sqrt{n} \sum_{k=1}^{n} (v_{k+1} + v_k)^2 - \frac{c}{\sqrt{n}} \sum_{k=1}^{n} k v_k^2, \tag{5.0.3}$$

where

$$Z_k = \sqrt{n}\big(c_{n-k+1}^2 s_{n-k+1}'^2 - \mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2]\big), \quad \widetilde{Z}_k = \sqrt{n}\big(s_{n-k+1}^2 c_{n-k}'^2 - \mathbb{E}[s_{n-k+1}^2]\mathbb{E}[c_{n-k}'^2]\big),$$
$$\text{and } Y_k = \sqrt{n}\left(c_{n-k}s_{n-k+1}c_{n-k}'s_{n-k}' - \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c_{n-k}'s_{n-k}']\right). \tag{5.0.4}$$

Then, there are constants $\alpha_1 > \alpha_2 > 0$ *such that* $J_{\alpha_1}(v) \le J(v) \le J_{\alpha_2}(v)$ *for all* $v \in \mathbb{R}^n$.

5.1 Lower Bound

The lower bound is much easier to obtain than the upper bound. After adding and subtracting the expectations of the random variables in 5.0.1, it suffices to show that for some $\alpha > 0$

$$\begin{aligned} - \quad & \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c_{n-k}'s_{n-k}'](v_{k+1}+v_k)^2 + \left(\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2] + \mathbb{E}[c_{n-k}'^2]\mathbb{E}[s_{n-k+1}^2]\right)v_k^2 \\ + \quad & \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c_{n-k}'s_{n-k}']\left(v_{k+1}^2 + v_k^2\right) - \gamma v_k^2 \end{aligned}$$

is bounded below by

$$-\alpha(v_{k+1}+v_k)^2 - \frac{\alpha k}{n}v_k^2 \tag{5.1.1}$$

for $1 \le k \le n$. By adding to the denominators and subtracting from the numerators, one easily gets the lower bound

$$\begin{aligned} \mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2] &= \frac{(a+1)n-k+1}{(a+b+2)n-2k+2} \cdot \frac{(a+b+1)n-k+2}{(a+b+2)n-2k+3} \\ &\ge \frac{(a+1)n-k}{(a+b+2)n} \cdot \frac{(a+b+1)n-k}{(a+b+2)n} \\ &= \frac{(a+1)(a+b+1)}{(a+b+2)^2}\left(1-\frac{k}{(a+1)n}\right)\left(1-\frac{k}{(a+b+1)n}\right). \end{aligned}$$

Because a and b are nonnegative one easily has

$$\begin{aligned} \mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2] &\ge \frac{(a+1)(a+b+1)}{(a+b+2)^2}\left(1-\frac{k}{n}\right)^2 \\ &= \frac{(a+1)(a+b+1)}{(a+b+2)^2}\left(1-\frac{2k}{n}+\left(\frac{k}{n}\right)^2\right) \\ &\ge \frac{(a+1)(a+b+1)}{(a+b+2)^2} - \frac{2(a+1)(a+b+1)}{(a+b+2)^2}\frac{k}{n}, \end{aligned} \tag{5.1.2}$$

where the positivity of $(k/n)^2$ allows it to be thrown away. A nearly identical argument can be used to produce the lower bound

$$\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2] \geq \frac{(b+1)}{(a+b+2)^2} - \frac{2(b+1)}{(a+b+2)^2}\frac{k}{n}. \tag{5.1.3}$$

Arguing in the same fashion will also produce a lower bound for $\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}]$, but before this is done the Chapter 4 bounds need to be used. By 4.2.4 and then using the fact that $\beta \geq 1$, one has

$$\begin{aligned}
\mathbb{E}[c_{n-k}] &\geq \sqrt{\frac{\beta[(a+1)n-k]-1}{\beta[(a+b+2)n-2k]}} \geq \sqrt{\frac{(a+1)n-k-1}{(a+b+2)n-2k}}, \\
\mathbb{E}[s_{n-k+1}] &\geq \sqrt{\frac{\beta[(b+1)n-k+1]-1}{\beta[(a+b+2)n-2k+2]}} \geq \sqrt{\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}},
\end{aligned}$$

and by 4.2.5, one also has

$$\begin{aligned}
\mathbb{E}[c'_{n-k}s'_{n-k}] &\geq \frac{\sqrt{\beta[n-k]-1}\sqrt{\beta[(a+b+1)n-k+1]-\frac{1}{2}}}{\beta[(a+b+2)n-2k+1]} \\
&\geq \sqrt{\frac{n-k-1}{(a+b+2)n-2k+1}\cdot\frac{(a+b+1)n-k}{(a+b+2)n-2k+1}}.
\end{aligned}$$

Combining these three bounds yields

$$\begin{aligned}
&\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}] \\
&\quad \geq \sqrt{\frac{(a+1)n-k-1}{(a+b+2)n-2k}\cdot\frac{(b+1)n-k}{(a+b+2)n-2k+2}\cdot\frac{n-k-1}{(a+b+2)n-2k+1}\cdot\frac{(a+b+1)n-k}{(a+b+2)n-2k+1}}.
\end{aligned}$$

Then, just as above, one has

$$\begin{aligned}
\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}] &\geq \frac{\sqrt{(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}\left(1-\frac{2(k+1)}{n}\right)^2 \\
&\geq \frac{\sqrt{(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}\left(1-\frac{4k}{n}\right)^2 \\
&\geq \frac{\sqrt{(a+1)(b+1)(a+b+1)}}{(a+b+2)^2} - \frac{2\sqrt{2(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}\frac{k}{n},
\end{aligned}$$

and so

$$\begin{aligned}
&\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}](v_{k+1}^2+v_k^2) && (5.1.4)\\
\geq\ & \left[\frac{\sqrt{(a+1)(b+1)(a+b+1)}}{(a+b+2)^2} - \frac{2\sqrt{2(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}\frac{k}{n}\right](v_{k+1}^2+v_k^2)\\
\geq\ & \frac{2\sqrt{(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}v_k^2 - \frac{2\sqrt{2(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}\left(\frac{k}{n}+\frac{k-1}{n}\right)v_k^2\\
\geq\ & \frac{2\sqrt{(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}v_k^2 - \frac{4\sqrt{2(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}\frac{k}{n}v_k^2. && (5.1.5)
\end{aligned}$$

Combining 5.1.2, 5.1.3, and 5.1.4 shows that

$$\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s'^2_{n-k+1}]\,v_k^2 + \mathbb{E}[c'^2_{n-k}]\mathbb{E}[s^2_{n-k+1}]\,v_k^2 + \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}]\,(v_{k+1}^2+v_k^2)$$

is bounded below by

$$\gamma^2 v_k^2 - \frac{\alpha k}{n}v_k^2, \qquad (5.1.6)$$

where γ is defined in 4.1.2 and

$$\alpha = \frac{2(a+1)(a+b+1)}{(a+b+2)^2} + \frac{2(b+1)}{(a+b+2)^2} + \frac{4\sqrt{2(a+1)(b+1)(a+b+1)}}{(a+b+2)^2}. \qquad (5.1.7)$$

To finish the proof of the lower bound simply notice that because $Beta(s,t)$ random variables are bounded above by 1, one has the trivial bound

$$-\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}](v_{k+1}+v_k)^2 \geq -(v_{k+1}+v_k)^2,$$

so redefining alpha to be the maximum of 1 and the constant in 5.1.7 completes the proof of the lower bound.

5.2 Upper Bound

The upper bound is the more difficult of the two. As in the previous subsection, it is enough to show to find some $\alpha > 0$ such that

$$\begin{aligned}
-\ & \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}](v_{k+1}+v_k)^2 + \left(\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s'^2_{n-k+1}] + \mathbb{E}[c'^2_{n-k}]\mathbb{E}[s^2_{n-k+1}]\right)v_k^2\\
+\ & \left(\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}]\right)\,(v_{k+1}^2+v_k^2) - \gamma v_k^2 && (5.2.1)
\end{aligned}$$

is bounded above by

$$-\alpha(v_{k+1}+v_k)^2-\frac{\alpha k}{n}v_k^2 \tag{5.2.2}$$

for $1\le k\le n$.

Claim 17. *For $a\ge 0$, $b>0$, and $1\le k\le n$*

$$\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s'^2_{n-k+1}]\,v_k^2+\mathbb{E}[c'^2_{n-k}]\mathbb{E}[s^2_{n-k+1}]\,v_k^2+\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}]\,(v_{k+1}^2+v_k^2)$$

is bounded above by

$$\left(\sqrt{xy}+\sqrt{(1-x)(1-y)}\right)^2\,v_k^2,$$

where

$$x=\frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\quad \text{and}\quad y=\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2}.$$

Remark 18. The proof of this claim is fairly elementary, but it is also quite tedious. For that reason, it will be done at the end of the section.

One also needs the following lemma, which will also be proved at the end of the section.

Lemma 19. *Let $x,y\in(0,1)$ with $y=x(k)+\delta(k)$, and $\delta(k)>0$. If δ is increasing in k and $|x'(k)|<\delta'(k)$, then the function*

$$f(x,k)=\sqrt{xy}+\sqrt{(1-x)(1-y)}$$

is decreasing as k increases.

Essentially, Lemma 19 says that f is increasing as the distance between x and y is decreases. Notice, that

$$\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+3}\ge\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2}\ge\frac{(a+1)n-k+1}{(a+b+2)n-2k+2},$$

and so by Lemma 19, $f(x,y)$ increases by letting

$$y=\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2}.$$

It is not hard to show that y is increasing and concave when $1 \le k \le n-1$, so by Lemma 19, y can be replaced with its tangent line

$$y = \frac{a+b+1}{a+b+2} + \frac{a+b}{(a+b+2)^2} \cdot \frac{k-1}{n}$$

to bound $f(x,y)$ above by

$$\left(\sqrt{\frac{a+1}{a+b+2}} \sqrt{\frac{a+b+1}{a+b+2} + \frac{a+b}{(a+b+2)^2} \cdot \frac{k-1}{n}} + \sqrt{\frac{b+1}{a+b+2}} \sqrt{\frac{1}{a+b+2} - \frac{a+b}{(a+b+2)^2} \cdot \frac{k-1}{n}} \right)^2 .$$

Using the fact that $\sqrt{1+x} \le 1 + \frac{x}{2}$, $f(x,y)$ is in turn bounded by

$$\left(\gamma - C\frac{k-1}{n} \right)^2$$

where

$$C = -\frac{1}{2} \frac{a+b}{(a+b+2)^2} \left(\sqrt{\frac{a+1}{a+b+2}} - \sqrt{b+1} \right).$$

Notice that

$$\sqrt{\frac{a+1}{a+b+2}} - \sqrt{b+1} \le 1 - \sqrt{b+1},$$

and thus $c > 0$. Then

$$\begin{aligned} f(x,y) &\le \left(\gamma - C\frac{k-1}{n} \right)^2 \\ &= \gamma^2 - 2\gamma C \frac{k-1}{n} + C^2 \left(\frac{k-1}{n} \right)^2, \end{aligned}$$

and is not hard to show that $2\gamma > C$. Therefore,

$$\mathbb{E}[c^2_{n-k+1}]\mathbb{E}[s'^2_{n-k+1}]\, v_k^2 + \mathbb{E}[c'^2_{n-k}]\mathbb{E}[s^2_{n-k+1}]\, v_k^2 + \left(\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}] \right) (v_{k+1}^2 + v_k^2)$$

is bounded by

$$f(x,y) \le \gamma^2 - 2\alpha \frac{k-1}{n} \le \gamma^2 - \frac{\alpha k}{n}$$

for some $\alpha > 0$.

Case 1 ($1 \le k \le \frac{n}{2} - 1$): Assume that $1 \le k \le n/2 - 1$. From the previous section,

$$\begin{aligned} &\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}] \\ &\quad \ge \sqrt{\frac{(a+1)n-k-1}{(a+b+2)n-2k} \cdot \frac{(b+1)n-k}{(a+b+2)n-2k+2} \cdot \frac{n-k-1}{(a+b+2)n-2k+1} \cdot \frac{(a+b+1)n-k}{(a+b+2)n-2k+1}}, \end{aligned}$$

and because $1 \le k \le \frac{n}{2} - 1 < \frac{n}{2}$, one easily has

$$\begin{aligned}
\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}] &\ge \sqrt{\frac{(a+\frac{1}{2})n}{(a+b+2)n} \cdot \frac{(b+\frac{1}{2})n}{(a+b+2)n} \cdot \frac{\frac{n}{2}}{(a+b+2)n} \cdot \frac{(a+b+\frac{1}{2})n}{(a+b+2)n}} \\
&\ge \frac{\sqrt{(2a+1)(2b+1)(2a+2b+1)}}{4(a+b+2)^2} \\
&\ge \frac{\sqrt{(a+1)(b+1)(a+b+1)}}{4(a+b+2)^2}
\end{aligned}$$

With $\alpha = \frac{1}{4(a+b+2)^2}$, this shows that

$$-\sum_{k=1}^{\lfloor \frac{n}{2} \rfloor - 1} \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}](v_{k+1}+v_k)^2 \le -\alpha \sum_{k=1}^{\lfloor \frac{n}{2} \rfloor - 1} (v_{k+1}+v_k)^2$$

for some $\alpha > 0$.

Case 2: Assume that $\frac{n}{2} \le k \le n$. This is the easier of the two cases because

$$-\sum_{k=\lfloor \frac{n}{2} \rfloor + 1}^{n} \mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}](v_{k+1}+v_k)^2$$

is negative, and so it can be completely thrown away. By Lemma 19, letting k take its smallest value, $k = \lfloor \frac{n}{2} \rfloor$ produces

$$\begin{aligned}
\left(\sqrt{xy} + \sqrt{(1-x)(1-y)}\right)^2 &\le \left(\frac{\sqrt{(a+1/2)(a+b+1/2)}}{a+b+1} + \frac{(b+1/2)(1/2)}{a+b+1}\right)^2 \\
&< \left(\frac{\sqrt{(a+1)(a+b+1)}}{a+b+2} + \frac{\sqrt{b+1}}{a+b+2}\right)^2.
\end{aligned}$$

Thus, the k^{th} term of 5.2.1 is bounded above by $-\alpha v_k^2$ where

$$\begin{aligned}
-\alpha &= \left(\frac{\sqrt{(a+1/2)(a+b+1/2)}}{a+b+1} + \frac{(b+1/2)(1/2)}{a+b+1}\right)^2 \\
&\quad - \left(\frac{\sqrt{(a+1)(a+b+1)}}{a+b+2} + \frac{\sqrt{b+1}}{a+b+2}\right)^2.
\end{aligned}$$

Notice that

$$\begin{aligned} -\alpha \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} v_k^2 &= -\frac{\alpha}{4} \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} \left(v_k^2 + v_k^2\right) - \frac{\alpha}{2} \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} v_k^2 \\ &\leq -\frac{\alpha}{4} \sum_{k=\lfloor \frac{n}{2} \rfloor}^{n} \left(v_{k+1}^2 + v_k^2\right) - \frac{\alpha}{2n} \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} k v_k^2 \\ &\leq -\frac{\alpha}{8} \sum_{k=\lfloor \frac{n}{2} \rfloor}^{n} \left(v_{k+1} + v_k\right)^2 - \frac{\alpha}{8n} \sum_{k=\lfloor \frac{n}{2} \rfloor+1}^{n} k v_k^2 \end{aligned}$$

where the last inequality came from the fact that

$$(s+t)^2 \leq 2\left(s^2+t^2\right).$$

Theorefore, there is some constant $\alpha > 0$ such that the desired bound holds on the range $\frac{n}{2} \leq k \leq n$. At this point, all that remains is to establish Claim 17 and Lemma 19.

Proof of Claim 17. The first step is to rewrite the given terms $\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2]$ and $\mathbb{E}[c_{n-k}'^2]\mathbb{E}[s_{n-k+1}^2]$ as

$$xy + error \qquad \text{and} \qquad (1-x)(1-y) + error,$$

where

$$x = \frac{(a+1)n - k + 1}{(a+b+2)n - 2k + 2} \quad \text{and} \quad y = \frac{(b+1)n - k + 1}{(a+b+2)n - 2k + 2}.$$

The total error will then be negative, and hence

$$\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2] + \mathbb{E}[c_{n-k}'^2]\mathbb{E}[s_{n-k+1}^2] \leq xy + (1-x)(1-y). \tag{5.2.3}$$

Towards this end, write

$$\begin{aligned}
\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2] &= \frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+3} \\
&= \frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2} \\
&\quad \cdot\left(1-\frac{1}{(a+b+2)n-2k+3}\right)\left(1+\frac{1}{(a+b+1)n-k+1}\right) \\
&= \frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2} \\
&\quad \cdot\left(1+\frac{n-k+1}{[(a+b+2)n-2k+3][(a+b+1)n-k+1]}\right) \\
&= \frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2} \\
&\quad +\frac{[(a+1)n-k+1][n-k+1]}{[(a+b+2)n-2k+2]^2[(a+b+2)n-2k+3]}. \qquad (5.2.4)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\mathbb{E}[c_{n-k}'^2]\mathbb{E}[s_{n-k+1}^2] &= \frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k}{(a+b+2)n-2k+1} \\
&= \frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k+1}{(a+b+2)n-2k+2} \\
&\quad -\frac{[(b+1)n-k+1][(a+b+1)n-k+1]}{[(a+b+2)n-2k+2]^2[(a+b+2)n-2k+1]}. \qquad (5.2.5)
\end{aligned}$$

Adding the error terms from 5.2.4 and 5.2.5 is bounded above by zero because

$$\begin{aligned}
(a+1)n-k+1 &\leq (a+b+1)n-k+1, \\
n-k+1 &\leq (b+1)n-k+1, \text{ and} \\
\frac{1}{(a+b+2)n-2k+3} &\leq \frac{1}{(a+b+2)n-2k+1},
\end{aligned}$$

and so 5.2.3 holds true.

The next step is to show that, for the same values of x and y,

$$\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c_{n-k}'s_{n-k}']\,(v_{k+1}^2+v_k^2) \leq \sqrt{x(1-x)y(1-y)}\,v_k^2. \qquad (5.2.6)$$

By Jensen's inequality, the square root can be taken outside the expectation to get

$$\mathbb{E}\sqrt{X} \leq \sqrt{\mathbb{E}X},$$

and so

$$\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}] \leq \sqrt{\mathbb{E}[c^2_{n-k}]\mathbb{E}[s^2_{n-k+1}]\mathbb{E}[c'^2_{n-k}s'^2_{n-k}]},$$

which is in turn equal to

$$\sqrt{\frac{(a+1)n-k}{(a+b+2)n-2k}\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k}{(a+b+2)n-2k+1}\frac{\frac{\beta}{2}[(a+b+1)n-k+1]}{\frac{\beta}{2}[(a+b+2)n-2k+1]+1}}.$$

Because $\beta > 0$, $\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}]$ is bounded above by

$$\sqrt{\frac{(a+1)n-k}{(a+b+2)n-2k}\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k}{(a+b+2)n-2k+1}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+1}},$$

and hence $\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c'_{n-k}s'_{n-k}](v^2_{k+1}+v^2_k)$ is bounded above by

$$\sqrt{\frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(b+1)n-k+2}{(a+b+2)n-2k+4}\frac{n-k+1}{(a+b+2)n-2k+3}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+3}}\,v_k^2$$
$$+\sqrt{\frac{(a+1)n-k}{(a+b+2)n-2k}\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k}{(a+b+2)n-2k+1}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+1}}\,v_k^2.$$

Continuing will require two trivial facts. These are, for $1 < u \leq v < \infty$,

$$\frac{1}{u^2} \leq \frac{1}{(u-1)(u+1)}, \tag{5.2.7}$$

and

$$u(v+1) \leq (u+1)v. \tag{5.2.8}$$

By 5.2.7

$$\frac{n-k}{(a+b+2)n-2k+1}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+1} \leq \frac{n-k}{(a+b+2)n-2k}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2},$$

and by 5.2.8

$$[(a+1)n-k][(a+b+1)n-k+1] \leq [(a+1)n-k+1][(a+b+1)n-k].$$

Together this means that

$$\sqrt{\frac{(a+1)n-k}{(a+b+2)n-2k}\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k}{(a+b+2)n-2k+1}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+1}}$$
$$\leq \sqrt{\frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k}{(a+b+2)n-2k}\frac{(a+b+1)n-k}{(a+b+2)n-2k}}.$$

Similar reasoning produces

$$\begin{aligned}
&\sqrt{\frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(b+1)n-k+2}{(a+b+2)n-2k+4}\frac{n-k+1}{(a+b+2)n-2k+3}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+3}} \\
\leq\ &\sqrt{\frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\frac{(b+1)n-k+1}{(a+b+2)n-2k+2}\frac{n-k+2}{(a+b+2)n-2k+4}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}},
\end{aligned}$$

and so one needs to show that

$$\begin{aligned}
&\sqrt{\frac{n-k}{(a+b+2)n-2k}\frac{(a+b+1)n-k}{(a+b+2)n-2k}}+\sqrt{\frac{n-k+2}{(a+b+2)n-2k+4}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}} \\
\leq\ & 2\cdot\sqrt{\frac{n-k+1}{(a+b+2)n-2k+2}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2}}.
\end{aligned}$$

Recall that, if f is concave, then for $t \in [0,1]$

$$(1-t)f(x)+tf(y) \leq f((1-t)x+ty).$$

With $f(x)=\sqrt{x}$ and $t=1/2$, one has

$$\begin{aligned}
&\sqrt{\frac{n-k}{(a+b+2)n-2k}\frac{(a+b+1)n-k}{(a+b+2)n-2k}}+\sqrt{\frac{n-k+2}{(a+b+2)n-2k+4}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}} \\
\leq\ & 2\cdot\sqrt{\frac{1}{2}\frac{n-k}{(a+b+2)n-2k}\frac{(a+b+1)n-k}{(a+b+2)n-2k}+\frac{1}{2}\frac{n-k+2}{(a+b+2)n-2k+4}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}}.
\end{aligned}$$

Apply the same logic with $f(x)=x(1-x)$ and $t=1/2$ to get

$$\begin{aligned}
&2\cdot\sqrt{\frac{1}{2}\frac{n-k}{(a+b+2)n-2k}\frac{(a+b+1)n-k}{(a+b+2)n-2k}+\frac{1}{2}\frac{n-k+2}{(a+b+2)n-2k+4}\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}} \\
&\sqrt{\left(\frac{n-k}{(a+b+2)n-2k}+\frac{n-k+2}{(a+b+2)n-2k+4}\right)\left(\frac{(a+b+1)n-k}{(a+b+2)n-2k}+\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}\right)}.
\end{aligned}$$

Fact 20. *Let $x(k) \in [0,1]$. Then $x(k)(1-x(k))$ increases as the distance between $x(k)$ and $1-x(k)$ decreases.*

Apply Fact 20 to see that

$$\begin{aligned}
&\frac{1}{2}\left(\frac{n-k}{(a+b+2)n-2k}+\frac{n-k+2}{(a+b+2)n-2k+4}\right)\cdot\frac{1}{2}\left(\frac{(a+b+1)n-k}{(a+b+2)n-2k}+\frac{(a+b+1)n-k+2}{(a+b+2)n-2k+4}\right) \\
\leq\ & \frac{n-k+1}{(a+b+2)n-2k+2}\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2},
\end{aligned}$$

and hence 5.2.6 holds. Combining 5.2.3 and 5.2.6 gives the upper bound

$$\mathbb{E}[c_{n-k+1}^2]\mathbb{E}[s_{n-k+1}'^2]\,v_k^2+\mathbb{E}[c_{n-k}'^2]\mathbb{E}[s_{n-k+1}^2]\,v_k^2+\mathbb{E}[c_{n-k}]\mathbb{E}[s_{n-k+1}]\mathbb{E}[c_{n-k}'s_{n-k}']\,(v_{k+1}^2+v_k^2)$$

is bounded above by

$$\left(\sqrt{xy}+\sqrt{(1-x)(1-y)}\right)^2\ v_k^2,$$

where

$$x=\frac{(a+1)n-k+1}{(a+b+2)n-2k+2}\quad\text{and}\quad y=\frac{(a+b+1)n-k+1}{(a+b+2)n-2k+2}.$$

□

Proof of Lemma 19. Again, differentiation in k yields

$$\begin{aligned}\frac{df}{dk} &= \frac{d}{dk}\left[\sqrt{x(k)(x(k)+\delta(k))}+\sqrt{(1-x(k))(1-x(k)-\delta(k))}\right]\\ &= \frac{1}{2}\left[\sqrt{1-\frac{\delta(k)}{x(k)+\delta(k)}}-\sqrt{1+\frac{\delta(k)}{1-(x(k)+\delta(k))}}\right]\cdot(x'(k)+\delta'(k))\\ &\quad+\frac{1}{2}\left[\sqrt{1+\frac{\delta(k)}{x(k)}}-\sqrt{1-\frac{\delta(k)}{1-x(k)}}\right]\cdot x'(k)\end{aligned}\tag{5.2.9}$$

Because $0<x,y<1$, one easily has

$$\sqrt{1-\frac{\delta(k)}{x(k)+\delta(k)}}-\sqrt{1+\frac{\delta(k)}{1-(x(k)+\delta(k))}}<0$$

and

$$\sqrt{1+\frac{\delta(k)}{x(k)}}-\sqrt{1-\frac{\delta(k)}{1-x(k)}}>0.$$

If $x'(k)<0$, then the proof is completed by noticing that $x'(k)+\delta'(k)>0$. If $x'(k)\geq 0$, then rewrite 5.2.9 as

$$\begin{aligned}&\frac{1}{2}\left[\sqrt{1+\frac{\delta(k)}{x(k)}}-\sqrt{1-\frac{\delta(k)}{1-x(k)}}+\sqrt{1-\frac{\delta(k)}{x(k)+\delta(k)}}-\sqrt{1+\frac{\delta(k)}{1-(x(k)+\delta(k))}}\right]\cdot x'(k)\\ &+\frac{1}{2}\left[\sqrt{1-\frac{\delta(k)}{x(k)+\delta(k)}}-\sqrt{1+\frac{\delta(k)}{1-(x(k)+\delta(k))}}\right]\cdot\delta'(k).\end{aligned}$$

Because $x(k)\leq x(k)+\delta(k)$

$$\sqrt{1+\frac{\delta(k)}{x(k)}}-\sqrt{1-\frac{\delta(k)}{1-x(k)}}<-\sqrt{1-\frac{\delta(k)}{x(k)+\delta(k)}}+\sqrt{1+\frac{\delta(k)}{1-(x(k)+\delta(k))}},$$

and so the first term is negative. The second term is clearly negative as well, so the proof is complete.

□

Chapter 6

Right-Tail Upper Bound

This chapter contains the proof of Theorem 13, the right-tail upper bound. The most important modification to the proof of the β-Hermite case, found in [16], is Lemma 24. The need for this modification is closely tied to the term containing $c'_{n-k}s'_{n-k}$. More specifically, the fact that the function $\sqrt{X(1-X)}$ is not $1-$Lipshitz.

Proposition 21. *Consider the model quadratic form,*

$$J_\alpha(v,z) = \sum_{k=1}^{n} z_k v_k^2 - \alpha\sqrt{n}\sum_{k=1}^{n}(v_{k+1}+v_k)^2 - \frac{\alpha}{\sqrt{n}}\sum_{k=1}^{n} k v_k^2, \tag{6.0.1}$$

for fixed $\alpha > 0$ and independent mean-zero random variables $\{z_k\}_{k=1,\ldots,n}$ satisfying $\mathbb{E}[e^{\lambda z_k}] \leq 2 \cdot e^{c\lambda^2/\beta(a+b)}$ for some $c > 0$ and all $\lambda \in \mathbb{R}$. There is a $C = C(\alpha, c)$ so that

$$\mathbb{P}\left(\sup_{||v||_2=1} J_\alpha(v,z) \geq \varepsilon\gamma\sqrt{n}\right) \leq \left(1 - e^{\beta(a+b)/C}\right)^{-1} e^{-\beta(a+b)\varepsilon^{3/2}n/C}$$

for all $\varepsilon \in (0,1]$ and $n \geq 1$.

The proof of this proposition will require the use of the following Lemma. Refer to [16] for a proof.

Lemma 22. *Let $s_1, s_2, \ldots, s_k, \ldots$ be real numbers, and let $S_k = \sum_{l=1}^{k} s_l, S_0 = 0$. Let further $t_1, \ldots, t_n$ be real numbers, $t_0 = t_{n+1} = 0$. Then, for every integer $m \geq 1$,*

$$\sum_{k=1}^{n} s_k t_k = \frac{1}{n}\sum_{k=1}^{n}[s_{k+m-1} - S_{k-1}]t_k + \sum_{k=0}^{n}\left(\frac{1}{m}\sum_{l=k}^{k+m-1}[s_l - s_k]\right)(t_{k+1} - t_k).$$

Proof of Proposition 21. Apply Lemma 22 with $s_k = z_k$ and $t_k = v_k^2$ to get

$$\begin{aligned}\sum_{k=1}^{n} z_k v_k^2 &\leq \frac{1}{m}\sum_{k=1}^{n} |S_{k+m-1} - S_{k-1}| v_k^2 + \sum_{k=0}^{n}\left(\frac{1}{m}\sum_{\ell=k}^{k+m-1} |S_\ell - S_k|\right) |v_{k+1}^2 - v_k^2| \\ &\leq \frac{1}{m}\sum_{k=1}^{n} \Delta_m(k-1) v_k^2 + \sum_{k=0}^{n} \Delta_m(k) |v_{k+1} + v_k||v_{k+1} - v_k|\end{aligned}$$

where

$$\Delta_m(k) = \max_{k+1\leq \ell \leq k+m} |S_\ell - S_k|, \quad \text{for } k = 0, \ldots, n. \tag{6.0.2}$$

By applying the Cauchy-Schwarz inequality

$$\sum_{k=1}^{n} z_k v_k^2 \leq \frac{1}{m}\sum_{k=1}^{n} \Delta_m(k-1) v_k^2 + \lambda \sum_{k=0}^{n} (v_{k+1} + v_k)^2 + \frac{1}{4\lambda}\sum_{k=0}^{n} \Delta_m(k)^2 (v_{k+1} - v_k)^2$$

for every $\lambda > 0$. With $\lambda = \alpha\sqrt{n}$, one has

$$\sup_{||v||_2=1} J_\alpha(z,v) \leq \max_{1\leq k\leq n}\left(\frac{1}{m}\,\Delta_m(k-1) + \frac{1}{2\alpha\sqrt{n}}\,\left[\Delta_m(k-1)^2 + \Delta_m(k)^2\right] - \alpha\frac{k}{\sqrt{n}}\right). \tag{6.0.3}$$

Now, if $(j-1)m + 1 \leq k \leq jm$, $1 \leq j \leq [n/m] + 1$, then the following inequality holds

$$\Delta_m(k) \vee \Delta_m(k-1) \leq 2\Delta_{2m}\big((j-1)m\big).$$

This implies that

$$\begin{aligned}&\sup_{||v||_2=1} J_\alpha(z,v) \\ &\leq \max_{1\leq j\leq [\frac{n}{m}]+1}\left(\frac{2}{m}\,\Delta_{2m}\big((j-1)m\big) + \frac{4}{\alpha\sqrt{n}}\,\Delta_{2m}\big((j-1)m\big)^2 - \alpha\,\frac{(j-1)m+1}{\sqrt{n}}\right).\end{aligned} \tag{6.0.4}$$

In order to continue, one must have a tail bound on $\Delta_{2m}(J)$ for any integer $J \geq 0$. Using Doob's maximal inequality and the assumptions on z_k, for every $\lambda > 0$ and $t > 0$, one has

$$\begin{aligned}\mathbb{P}\Big(\max_{1\leq \ell\leq 2m} S_\ell \geq t\Big) &\leq e^{-\lambda t}\,\mathbb{E}\left[e^{\lambda S_{2m}}\right] \\ &\leq e^{-\lambda t + \frac{2cm\lambda^2}{\beta(a+b)}}.\end{aligned}$$

Optimizing in λ, and then applying the same reasoning to the sequence $-S_\ell$ produces

$$\mathbb{P}\Big(\max_{1\leq \ell\leq 2m} |S_\ell| \geq t\Big) \leq 2\,e^{-\beta(a+b)t^2/8cm}.$$

Hence,

$$\mathbb{P}\Big(\Delta_{2m}(J) \geq t\Big) \leq 2\, e^{-\beta(a+b)t^2/8cm}, \tag{6.0.5}$$

for all integers $m \geq 1$ and $J \geq 0$, and every $t > 0$. One now has

$$\begin{aligned}
& \mathbb{P}\left(\max_{1\leq j\leq [n/m]+1}\left(\frac{2}{m}\,\Delta_{2m}\left((j-1)m\right) - \alpha\,\frac{(j-1)m+1}{2\sqrt{n}}\right) \geq \frac{\varepsilon\gamma\sqrt{n}}{2}\right) \\
\leq\ & \sum_{j=1}^{[n/m]+1} \mathbb{P}\left(\Delta_{2m}\left((j-1)m\right) \geq \frac{m}{2}\left[\alpha\,\frac{(j-1)m+1}{2\sqrt{n}} + \frac{\varepsilon\gamma\sqrt{n}}{2}\right]\right) \\
\leq\ & \sum_{j=1}^{[n/m]+1} 2\cdot\exp\left\{-\frac{\beta(a+b)m}{32c}\left[\alpha\,\frac{(j-1)m+1}{2\sqrt{n}} + \frac{\varepsilon\gamma\sqrt{n}}{2}\right]^2\right\} \\
=\ & \sum_{j=1}^{[n/m]+1} 2\cdot\exp\left\{-\frac{\beta(a+b)m}{32c}\left[\alpha^2\,\frac{[(j-1)m+1]^2}{4n} + \frac{\alpha\varepsilon\gamma[(j-1)m+1]}{2} + \frac{\varepsilon^2\gamma^2 n}{4}\right]\right\} \\
\leq\ & e^{-\beta(a+b)m\varepsilon^2\gamma^2 n/128c}\sum_{j=1}^{[n/m]+1} 2\cdot e^{-\beta(a+b)m\alpha\varepsilon\gamma[(j-1)m+1]/64c} \\
\leq\ & e^{-\beta(a+b)m\varepsilon^2\gamma^2 n/128c}\sum_{j=0}^{\infty} 2\cdot\left(e^{-\beta(a+b)m^2\alpha\varepsilon\gamma/64c}\right)^j \\
\leq\ & e^{-\beta(a+b)m\varepsilon^2\gamma^2 n/128c}\left(\frac{2}{1-e^{-\beta(a+b)m^2\alpha\varepsilon\gamma/64c}}\right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
& \mathbb{P}\left(\max_{1\leq j\leq [n/m]+1}\left(\frac{4}{\alpha\sqrt{n}}\,\Delta_{2m}\big((j-1)m\big)^2 - \alpha\,\frac{(j-1)m+1}{2\sqrt{n}}\right) \geq \frac{\varepsilon\gamma\sqrt{n}}{2}\right) \\
\leq\ & \sum_{j=1}^{[n/m]+1} \mathbb{P}\left(\Delta_{2m}\left((j-1)m\right)^2 \geq \frac{\alpha\sqrt{n}}{4}\left[\alpha\,\frac{(j-1)m+1}{2\sqrt{n}} + \varepsilon\gamma\sqrt{n}\right]\right) \\
\leq\ & \sum_{j=1}^{[n/m]+1} \mathbb{P}\left(\Delta_{2m}\left((j-1)m\right) \geq \left(\frac{\alpha\sqrt{n}}{4}\left[\alpha\,\frac{(j-1)m+1}{2\sqrt{n}} + \varepsilon\gamma\sqrt{n}\right]\right)^{1/2}\right) \\
\leq\ & \sum_{j=1}^{[n/m]+1} 2\cdot\exp\left\{-\beta(a+b)\frac{\alpha\sqrt{n}}{32cm}\left[\alpha\,\frac{(j-1)m+1}{2\sqrt{n}} + \frac{\varepsilon\gamma\sqrt{n}}{2}\right]\right\} \\
\leq\ & e^{-\beta(a+b)\alpha\varepsilon\gamma n/64cm}\sum_{j=0}^{\infty} 2\cdot\left(e^{-\beta(a+b)\alpha^2/64c}\right)^j \\
\leq\ & e^{-\beta(a+b)\alpha\varepsilon\gamma n/64cm}\left(\frac{2}{1-e^{-\beta(a+b)\alpha^2/64c}}\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}&\mathbb{P}\left(\sup_{||v||_2=1} J_\alpha(z,v) \geq \varepsilon\gamma\sqrt{n}\right)\\ &\leq \quad e^{-\beta(a+b)m\varepsilon^2\gamma^2 n/128c}\left(\frac{2}{1-e^{-\beta(a+b)m^2\alpha\varepsilon\gamma/64c}}\right) + e^{-\beta(a+b)\alpha\varepsilon\gamma n/64cm}\left(\frac{2}{1-e^{-\beta(a+b)\alpha^2/64c}}\right).\end{aligned}$$

Letting $m = [\varepsilon^{-1/2}]$ yields

$$\begin{aligned}&\mathbb{P}\left(\sup_{||v||_2=1} J_\alpha(z,v) \geq \varepsilon\gamma\sqrt{n}\right)\\ &\leq \quad e^{-\beta(a+b)\varepsilon^{3/2}\gamma^2 n/128c}\left(\frac{2}{1-e^{-\beta(a+b)\alpha\gamma/64c}}\right) + e^{-\beta(a+b)\alpha\varepsilon^{3/2}\gamma n/64c}\left(\frac{2}{1-e^{-\beta(a+b)\alpha^2/64c}}\right)\\ &\leq \quad e^{-\beta(a+b)\varepsilon^{3/2}n/C}\left(\frac{1}{1-e^{\beta(a+b)/C}}\right),\end{aligned}$$

which completes the proof of Proposition 21. □

Remark 23. When $\varepsilon > 1$, letting $m = 1$ means that the second term is larger, and so Theorem 13 becomes

$$\mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \geq \gamma\sqrt{n}(1+\varepsilon)\right) \leq C_\beta e^{-\beta(a+b)n\varepsilon/C_\beta}.$$

This will be used to prove Corollary 15.

Now, with Proposition 21 proved, the proof of Theorem 13 will begin to take shape. Notice that

$$\mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \geq \gamma\sqrt{n}(1+\varepsilon)\right) \quad \leq \quad \mathbb{P}\left(\sup_{||v||_2=1} J_\alpha(v) \geq \gamma\sqrt{n}\varepsilon\right).$$

Write 5.0.3 as

$$J_\alpha(v) = J_{\alpha/3}(Z_k, v) + J_{\alpha/3}(\widetilde{Z}_k, v) + \widetilde{J}_{\alpha/3}(Y_k, v) \tag{6.0.6}$$

where

$$\widetilde{J}_\alpha(Y_k, v) := 2\sum_{k=1}^{n-1} Y_k v_k v_{k+1} - \alpha\sqrt{n}\sum_{k=1}^{n}(v_{k+1}+v_k)^2 - \frac{\alpha}{\sqrt{n}}\sum_{k=1}^{n} k v_k^2. \tag{6.0.7}$$

The third term on the right is not quite ready for us to apply Proposition 21, but applying Lemma 22 to $\widetilde{J}_{\alpha/3}(Y_k, v)$ with $s_k = Y_k$ and $t_k = v_k v_{k+1}$, changes 6.0.4 to

$$\begin{aligned}&\sup_{||v||_2=1} J_\alpha(z,v)\\ &\leq \quad \max_{1\leq j\leq[\frac{n}{m}]+1}\left(\frac{4}{m}\,\Delta_{2m}\big((j-1)m\big) + \frac{8}{\alpha\sqrt{n}}\,\Delta_{2m}\big((j-1)m\big)^2 - \alpha\,\frac{(j-1)m+1}{\sqrt{n}}\right).\end{aligned}$$

The additional factor of two will be absorbed into the constant at the end.

In order to use Proposition 21 one $J_{\alpha/3}(Z_k, v)$, $J_{\alpha/3}(\widetilde{Z}_k, v)$, and $\widetilde{J}_{\alpha/3}(Y_k, v)$, all that remains is to establish the desired moment generating function bounds for Z_k, $\widetilde{Z}_k$, and Y_k . The following lemma produces these bounds.

Lemma 24. *Let X be a Beta(s,t) random variable with $s,t \geq 1$. Let $F(X)$ be a mean zero function such that $x(1-x)(F'(x))^2 \leq 1$ for $x \in [0,1]$. Then*

$$\mathbb{E}[e^{\lambda F(X))}] \leq e^{4\lambda^2/(s+t)}. \tag{6.0.8}$$

for every $\lambda \in \mathbb{R}$.

Remark 25. Ultimately, this proof will follow a modified Herbst argument. The Herbst argument requires a logarithmic-Sobolev inequality for the measure in question. For the Beta distributions in this problem, this comes from the fact that the Beta density is log-concave. If $F(X)$ is 1-Lipschitz, then the Herbst argument would complete the proof. However, $F(x) = \sqrt{X(1-X)}$ is not 1-Lipschitz, and so the log-Sobolev must be modified and F must satisfy the hypothesis stated in the Lemma.

Proof of Lemma 24. For $x \in [0,1]$, consider the Jacobi operator, $\mathcal{J}$, and the measure, μ, defined by

$$\mathcal{J} := \frac{1}{2}x(1-x)\frac{d^2}{dx^2} + \frac{1}{2}\left[s(1-x) - tx\right]\frac{d}{dx} \tag{6.0.9}$$

$$\mu_{s,t}(dx) := \frac{1}{Z_{s,t}}x^{s-1}(1-x)^{t-1}dx. \tag{6.0.10}$$

The carré du champ, $\Gamma_1(f,f) := \frac{1}{2}\left\{\mathcal{J}(f^2) - 2f\mathcal{J}(f)\right\}$, of the Jacobi operator is easily calculated to be

$$\Gamma_1(f,f) = \frac{1}{2}x(1-x)(f'(x))^2. \tag{6.0.11}$$

The next step is to use the Bakry-Emery Γ_2 criterion. The Bakry-Emery condition is satisfied with constant $c > 0$ if

$$\Gamma_2(f,f) = \frac{1}{2}\left\{\mathcal{J}\,\Gamma_1(f,f) - 2\,\Gamma_1(f,\mathcal{J}\,f)\right\} \geq \frac{1}{c}\Gamma_1(f,f). \tag{6.0.12}$$

For more discussion on Logarithmic Sobolev Inequalities and the Bakry-Emery condition see [9].

The calculation is straightforward, but it is also a bit tedious. One has

$$\begin{aligned}
\mathcal{J}\,\Gamma_1(f,f) &= -\frac{1}{2}x(1-x)(f'(x))^2 + \frac{1}{2}x^2(1-x)^2(f''(x))^2 \\
&\quad + \frac{1}{4}[s(1-x)-tx](1-2x)(f'(x))^2 + (1-2x)x(1-x)f'(x)f''(x) \\
&\quad + \frac{1}{2}x^2(1-x)^2 f'(x)f'''(x) + \frac{1}{2}[s(1-x)-tx]x(1-x)f'(x)f''(x)
\end{aligned}$$

and

$$\begin{aligned}
2\,\Gamma_1(f,\mathcal{J}\,f) &= -\frac{s+t}{2}x(1-x)(f'(x))^2 + \frac{1}{2}(1-2x)x(1-x)f'(x)f''(x) \\
&\quad + \frac{1}{2}x^2(1-x)^2 f'(x)f'''(x) + \frac{1}{2}[s(1-x)-tx]x(1-x)f'(x)f''(x).
\end{aligned}$$

Combining the above two equations produces

$$\begin{aligned}
\Gamma_2(f,f) &= \frac{s+t}{2}x(1-x)(f'(x))^2 + \frac{1}{4}x^2(1-x)^2(f''(x))^2 \\
&\quad + \frac{1}{4}[s(1-x)-tx](1-2x)(f'(x))^2 - \frac{1}{2}x(1-x)(f'(x))^2 - \frac{1}{4}(1-2x)^2(f'(x))^2 \\
&\quad + \frac{1}{4}(1-2x)^2(f'(x))^2 + \frac{1}{2}(1-2x)x(1-x)f'(x)f''(x) + \frac{1}{4}x^2(1-x)^2(f''(x))^2 \\
&= \frac{s+t}{2}x(1-x)(f'(x))^2 + \frac{1}{4}x^2(1-x)^2(f''(x))^2 \\
&\quad + \left[\frac{1}{4}[(s-1)(1-x)-(t-1)x](1-2x) - \frac{1}{2}x(1-x)\right](f'(x))^2 \\
&\quad + \frac{1}{4}\left[(1-2x)f'(x) + x(1-x)f''(x)\right]^2 \\
&= \frac{s+t}{2}x(1-x)(f'(x))^2 + \frac{1}{4}x^2(1-x)^2(f''(x))^2 \\
&\quad + \left[\frac{s-1}{4}(1-x)^2 + \frac{t-1}{4}x^2 - \frac{s+t}{4}x(1-x)\right](f'(x))^2 \\
&\quad + \frac{1}{4}\left[(1-2x)f'(x) + x(1-x)f''(x)\right]^2 \\
&\geq \frac{s+t}{4}x(1-x)(f'(x))^2.
\end{aligned}$$

The very last step used the hypothesis that $s,t \geq 1$. Thus, the Bakry-Emery condition is satisfied with constant $c = 4/(s+t)$.

Now, in order to use the Bakry-Emery condition one must first check that μ is the invariant measure of the Jacobi operator, which amounts to showing that $\int \mathcal{J}(f)d\mu = 0$. This is easily shown by integrating the first term in the following term by parts.

$$\int \mathcal{J}(f)d\mu \quad = \quad \frac{1}{Z_{s,t}}\int_0^1 x^s(1-x)^t f''(x) + [s(1-x) - tx]x^{s-1}(1-x)^{t-1}f'(x)\,dx = 0$$

and thus, μ is the invariant measure of the Jacobi operator.

Theorem 16 of [9] implies that if f is a positive, bounded, and continuous function with $\int f d\mu = 1$, then

$$\int f\log f d\mu \le 2c\int \Gamma_1(f^{1/2}, f^{1/2})d\mu. \tag{6.0.13}$$

This means that if f is bounded and continuous, then the inequality above will hold for the function $f^2/\|f\|_2$. Recalling 6.0.11, one has

$$\begin{aligned}\frac{1}{\|f\|_2}\int f^2\log\frac{f^2}{\|f\|_2}d\mu \quad &\le \quad 2c\int \Gamma_1\left(\frac{f}{\|f\|_2}, \frac{f}{\|f\|_2}\right)d\mu \\ &\le \quad \frac{c}{\|f\|_2}\int x(1-x)(f')^2 d\mu,\end{aligned}$$

and hence μ satisfies the following log-Sobolev inequality

$$\begin{aligned}\mathrm{Ent}_\mu(f^2) \quad &:= \quad \int f^2\log f^2 d\mu - \int f^2 d\mu \log\int f^2 d\mu \\ &= \quad \int f^2 \log\frac{f^2}{\|f\|_2}d\mu \\ &\le \quad c\int x(1-x)(f')^2 d\mu. \end{aligned} \tag{6.0.14}$$

The following is a modification to the Herbst argument, which is presented in [12]. First, apply 6.0.14 to the function $f^2 = e^{\lambda F - c\lambda^2/2}$, and then use the hypothesis on F to get

$$\begin{aligned}\int x(1-x)(f')^2 d\mu \quad &= \quad \frac{\lambda^2}{4}\int x(1-x)(F'(x))^2 e^{\lambda F(x) - c\lambda^2/2}d\mu \\ &\le \quad \frac{\lambda^2}{4}\int e^{\lambda F(x)-c\lambda^2/2}d\mu.\end{aligned}$$

From here on out we may follow [12] exactly to conclude that

$$\int e^{\lambda F}d\mu \le e^{c\lambda^2/2}$$

for all $\lambda \in \mathbb{R}$. □

Lemma 26. *Let X_1, X_2 be independent random variables such that $0 \le X_1, X_2 \le 1$. Assume that*

$$\mathbb{E}\left[e^{\lambda(X_i - \mathbb{E}X_i)}\right] \le e^{\lambda^2/c} \tag{6.0.15}$$

for some $c > 0$. Then

$$\mathbb{E}\left[e^{\lambda(X_1X_2 - \mathbb{E}X_1\mathbb{E}X_2)}\right] \le 2e^{\lambda^2/c}$$

for all $\lambda \in \mathbb{R}$.

Proof. By adding and subtracting $X_1\mathbb{E}[X_2]$ one has

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda(X_1X_2 - \mathbb{E}X_1\mathbb{E}X_2)}\right] &= \mathbb{E}\left[e^{\lambda X_1(X_2 - \mathbb{E}[X_2])}e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])}\right] \\
&= \mathbb{E}\left[e^{\lambda X_1(X_2 - \mathbb{E}[X_2])}e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])} \cdot \mathbb{1}_{\{X_2 - \mathbb{E}[X_2] \ge 0\}}\right] \\
&\quad + \mathbb{E}\left[e^{\lambda X_1(X_2 - \mathbb{E}[X_2])}e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])} \cdot \mathbb{1}_{\{X_2 - \mathbb{E}[X_2] < 0\}}\right].
\end{aligned}$$

If $\lambda > 0$ the fact that $0 \le X_1 \le 1$ implies that

$$e^{\lambda X_1(X_2 - \mathbb{E}[X_2])} \cdot \mathbb{1}_{\{X_2 - \mathbb{E}[X_2] \ge 0\}} \le e^{\lambda(X_2 - \mathbb{E}[X_2])},$$

and

$$e^{\lambda X_1(X_2 - \mathbb{E}[X_2])} \cdot \mathbb{1}_{\{X_2 - \mathbb{E}[X_2] < 0\}} \le 1.$$

Using the independence of X_1 and X_2, distribute the expectation to get

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda(X_1X_2 - \mathbb{E}X_1\mathbb{E}X_2)}\right] &\le \mathbb{E}\left[e^{\lambda(X_2 - \mathbb{E}[X_2])}e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])}\right] + \mathbb{E}\left[e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])}\right] \\
&= \mathbb{E}\left[e^{\lambda(X_2 - \mathbb{E}[X_2])}\right] \cdot \mathbb{E}\left[e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])}\right] + \mathbb{E}\left[e^{\lambda\mathbb{E}[X_2](X_1 - \mathbb{E}[X_1])}\right].
\end{aligned}$$

By 6.0.15 and the fact that $0 < X_i < 1$, one has

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda(X_1X_2 - \mathbb{E}[X_1X_2])}\right] &\le e^{\lambda^2/c} \cdot e^{\lambda^2/c} + e^{\lambda^2/c} \\
&\le e^{2\lambda^2/c} + e^{\lambda^2/c} \\
&\le 2 \cdot e^{2\lambda^2/c}.
\end{aligned}$$

The proof is essentially the same for $\lambda < 0$. If $\lambda = 0$, the proof is trivial.

$\square$

To finish the proof, let X is a $Beta(s,t)$ random variable. Then both $X - \mathbb{E}X$ and $\sqrt{X(1-X)} - \mathbb{E}\sqrt{X(1-X)}$ satisfy the Lemma 24 hypothesis on $F(X)$. Hence, by Lemma 24 one has

$$\mathbb{E}e^{\lambda(X-\mathbb{E}X)} \le e^{4\lambda^2/(s+t)} \qquad \text{and} \qquad \mathbb{E}e^{\lambda\left(\sqrt{X(1-X)}-\mathbb{E}\sqrt{X(1-X)}\right)} \le e^{4\lambda^2/(s+t)}.$$

Notice that the $Beta(s,t)$ random variables that make up Z_k, $\widetilde{Z}_k$, and Y_k all satisfy

$$s + t \ge \frac{\beta(a+b)n}{2},$$

and so Lemma 26 yields

$$\mathbb{E}e^{\lambda Z_k} \le 2e^{8\lambda^2/\beta(a+b)n} \qquad \text{and} \qquad \mathbb{E}e^{\lambda \widetilde{Z}_k} \le 2e^{8\lambda^2/\beta(a+b)n}.$$

Also, Lemma 26 can easily be extended to show that

$$\mathbb{E}e^{\lambda Y_k} \le 3e^{8\lambda^2/\beta(a+b)n}, \tag{6.0.16}$$

which completes the proof of Theorem 13.

Chapter 7

Left-Tail Upper Bound

By Lemma 16 there exists a constant $\alpha > 0$ such that

$$\begin{aligned}\mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \leq \gamma\sqrt{n}(1-\varepsilon)\right) &= \mathbb{P}\left(\sup_{\|v\|_2=1} J(v) \leq -\gamma\varepsilon\sqrt{n}\right) \\ &\leq \mathbb{P}\left(\sup_{\|v\|_2=1} J_\alpha(v) \leq -\gamma\varepsilon\sqrt{n}\right).\end{aligned}$$

One nice thing about the supremum here is that finding an upper bound can be done by choosing any test vector, $v \in \mathbb{R}^n$, such that $\|v\|_2 = 1$. If one normalizes by $\|v\|_2^2$, then the problem amounts to finding a bound for

$$\mathbb{P}\left(J_\alpha(v) \leq -C\gamma\varepsilon\sqrt{n}\,\|v\|_2^2\right)$$

where v can be any test vector in $\mathbb{R}^n$.

Fortunately, the choice of v can be the same as the choice for the β-Hermite ensemble in [16]. Namely, for $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ set

$$v_k = \begin{cases} \frac{k}{\varepsilon n} \wedge \left(1 - \frac{k}{\varepsilon n}\right) & \text{for } 1 \leq k \leq \varepsilon n \\ 0 & \text{otherwise.} \end{cases} \tag{7.0.1}$$

Borrowing the notation from [16], set

$$\|v\|_2^2 = \textstyle\sum_{k=1}^n v_k^2, \quad \|\nabla v\|_2^2 = \sum_{k=1}^n (v_{k+1} + v_k)^2, \text{ and } \quad \|\sqrt{k}v\|_2^2 = \sum_{k=1}^n kv_k^2.$$

For this choice of v, one easily has the following bounds

$$\|v\|_2^2 \sim \|v\|_4^4 \sim \varepsilon n, \quad \|\nabla v\|_2^2 \sim \tfrac{1}{\varepsilon n}, \text{ and } \quad \|\sqrt{k}v\|_2^2 \sim \varepsilon^2 n^2, \tag{7.0.2}$$

where $x \sim y$ indicates that there exists constants $c_2 \geq c_1 > 0$ such that $c_1 y \leq x \leq c_2 y$.

If $\varepsilon^{3/2} n \geq 1$, then

$$\begin{aligned}\mathbb{P}\left(J_\alpha(v,z) \leq -C\gamma\varepsilon\sqrt{n}\,\|v\|_2^2\right) &= \mathbb{P}\left(\sum_{k=1}^{n}(-z_k)v_k^2 \geq C\gamma\varepsilon\sqrt{n}\,\|v\|_2^2 - \alpha\sqrt{n}\|\nabla v\|_2^2 - \frac{\alpha}{\sqrt{n}}\|\sqrt{k}v\|_2^2\right) \\ &\leq \mathbb{P}\left(\sum_{k=1}^{n}(-z_k)v_k^2 \geq C\varepsilon^2 n^{3/2}\right). \qquad (7.0.3)\end{aligned}$$

The requirement on $\varepsilon^{3/2}n$ is needed to make sure that the $\|\nabla v\|_2^2$ term does not dominate the right hand side of this probability. Though it will not matter after the next step, it is possible for C to be negative at this point.

Now, for z_k satisfying the hypotheses in Proposition 21 one has

$$\begin{aligned}\mathbb{P}\left(\sum_{k=1}^{n}(-z_k)v_k^2 \geq t\right) &\leq e^{-\lambda t}\prod_{k=1}^{n}\mathbb{E}e^{-\lambda v_k^2 z_k} \\ &\leq e^{-\lambda t}\prod_{k=1}^{n}e^{c\lambda^2 v_k^4/\beta(a+b)} \\ &= e^{-\lambda t + \frac{c\|v\|_4^4\lambda^2}{\beta(a+b)}}\end{aligned}$$

for all $\lambda \in \mathbb{R}$. As in the proof of Proposition 21, optimize in λ to get

$$\mathbb{P}\left(\sum_{k=1}^{n}(-z_k)v_k^2 \geq t\right) \leq e^{-\frac{\beta(a+b)t^2}{4c\varepsilon n}}$$

where we used 7.0.2. Thus, for $t = C\varepsilon^2 n^{3/2}$ there exists a constant $C > 0$ such that

$$\mathbb{P}\left(\sum_{k=1}^{n}(-z_k)v_k^2 \geq C\varepsilon^2 n^{3/2}\right) \leq e^{-C\beta(a+b)\varepsilon^3 n^2}. \qquad (7.0.4)$$

Recall the definitions of Z_k, $\widetilde{Z}_k$, and Y_k given in 5.0.4. One has

$$J_\alpha(v) = J_{\alpha/3}(v,Z) + J_{\alpha/3}(v,\widetilde{Z}) + \widetilde{J}_{\alpha/3}(v,Y).$$

Since Z_k, $\widetilde{Z}_k$, and Y_k satisfy the hypotheses of Proposition 21, one has

$$\mathbb{P}\left(J_{\alpha/3}(v,Z) \leq -C\gamma\varepsilon\sqrt{n}\,\|v\|_2^2/3\right) \leq e^{-C_1\beta(a+b)\varepsilon^3 n^2},$$

$$\mathbb{P}\left(J_{\alpha/3}(v,\widetilde{Z}) \leq -C\gamma\varepsilon\sqrt{n}\,\|v\|_2^2/3\right) \leq e^{-C_2\beta(a+b)\varepsilon^3 n^2}.$$

The $\widetilde{J}$ term requires slightly more work, but the Caucy-Schwarz inequality, when applied to $\sum_{k=1}^{n-1} Y_k v_k v_{k+1}$, yields the bound

$$\mathbb{P}\left(\widetilde{J}_{\alpha/3}(v, Y) \leq -C\gamma\varepsilon\sqrt{n}\,\|v\|_2^2/3\right) \leq 2e^{-C_3\beta(a+b)\varepsilon^3 n^2}.$$

This completes the proof of Theorem 14.

Chapter 8

Variance Bound

Now that Theorems 13 and 14 have been established, a finite n bound on the variance of the largest eigenvalue follows quickly. Recall that in order to prove Corollary 15, one needs to show that

$$\mathrm{Var}\left[\lambda_{\max}(J_{\beta,n})\right] \leq C_\beta n^{-1/3}$$

for all $n \geq 1$.

The first step is to show that

$$\begin{aligned} \mathrm{Var}\left[\lambda_{\max}(J_{\beta,n})\right] &\leq \mathbb{E}[\left(\lambda_{\max}(J_{\beta,n}) - \gamma\sqrt{n}\right)^2] \\ &\leq \gamma^2 n \int_0^\infty \mathbb{P}\left(\left|\lambda_{\max}(J_{\beta,n}) - \gamma\sqrt{n}\right| \geq \gamma\sqrt{n}\varepsilon\right) d\varepsilon^2 \end{aligned}$$

The first inequality is because $\mathbb{E}[(X-C)^2]$ is minimized when $C = \mathbb{E}X$, and the second follows from applying Fubini's Theorem.

On the regime when $\lambda_{\max}(J_{\beta,n}) \leq \gamma\sqrt{n}$, use the fact that the Jacobi tridiagonal is a positive definite matrix (see 4.1.1). This means that all the eigenvalues are positive, and so when $\varepsilon > 1$ one has

$$\mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \leq \gamma\sqrt{n}(1-\varepsilon)\right) = 0. \tag{8.0.1}$$

If $\varepsilon \in (0,1]$, use Theorem 14 to say

$$\begin{aligned} \int_0^1 \mathbb{P}\left(\lambda_{\max}\left(J_\beta\right) \leq \gamma\sqrt{n}(1-\varepsilon)\right) d\varepsilon^2 &\leq \int_0^1 C_\beta e^{-\beta(a+b)\varepsilon^3 n^2/C_\beta} d\varepsilon^2 \\ &\leq C_\beta n^{-4/3} \int_0^\infty u^{-1/3} e^{-u} du \\ &\leq C_\beta n^{-4/3}, \end{aligned} \tag{8.0.2}$$

where C_β is a numerical constant. Combining 8.0.1 and 8.0.2 yields

$$\int_0^\infty \mathbb{P}\left(\lambda_{\max}(J_{\beta,n}) - \gamma\sqrt{n} \leq -\gamma\sqrt{n}\varepsilon\right) d\varepsilon^2 \leq C_\beta n^{-4/3}$$

for some C_β.

On the other regime, $\lambda_{\max}(J_{\beta,n}) \geq \gamma\sqrt{n}$, Theorem 13 can be used for $\varepsilon \in (0,1]$ to say

$$\begin{aligned}\int_0^1 \mathbb{P}\left(\lambda_{\max}(J_\beta) \geq \gamma\sqrt{n}(1+\varepsilon)\right) &\leq \int_0^1 C_\beta e^{-\beta(a+b)n\varepsilon^{3/2}/C_\beta} \\ &\leq C_\beta n^{-4/3}\int_0^\infty u^{1/3}e^{-u}du \\ &\leq C_\beta n^{-4/3}, \qquad (8.0.3)\end{aligned}$$

where C_β is a numerical constant. Now, if $\varepsilon > 1$ the proof of Theorem 13 can be tweaked slightly (see Remark 23) to produce

$$\mathbb{P}\left(\lambda_{\max}(J_\beta) \geq \gamma\sqrt{n}(1+\varepsilon)\right) \leq C_\beta e^{-\beta(a+b)n\varepsilon/C_\beta}.$$

Thus,

$$\begin{aligned}\int_1^\infty \mathbb{P}\left(\lambda_{\max}(J_\beta) \geq \gamma\sqrt{n}(1+\varepsilon)\right) d\varepsilon^2 &\leq \int_1^\infty C_\beta e^{-\beta(a+b)n\varepsilon/C_\beta} d\varepsilon^2 \\ &\leq C_\beta n^{-2}\int_0^\infty ue^{-u}du \\ &\leq C_\beta n^{-4/3}. \qquad (8.0.4)\end{aligned}$$

Combining 8.0.3 and 8.0.4 produces

$$\int_0^\infty \mathbb{P}\left(|\lambda_{\max}(J_{\beta,n}) - \gamma\sqrt{n}| \geq \gamma\sqrt{n}\varepsilon\right) d\varepsilon^2 \leq C_\beta n^{-4/3},$$

which completes the proof of Corollary 15.

Bibliography

[1] G.W. Anderson, A. Guionnet, and O. Zeitouni. An introduction to random matrices, volume 200. Cambridge University Press Cambridge, 2010.

[2] G.B. Arous, A. Dembo, and A. Guionnet. Aging of spherical spin glasses. Probability theory and related fields, 120(1):1–67, 2001.

[3] J. Baik, P. Deift, and K. Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. Journal of the American Mathematical Society, 12(4):1119–1178, 1999.

[4] B. Collins. Product of random projections, jacobi ensembles and universality problems arising from free probability. Probability theory and related fields, 133(3):315–344, 2005.

[5] I. Dumitriu and A. Edelman. Matrix models for beta ensembles. Journal of Mathematical Physics, 43:5830, 2002.

[6] I. Dumitriu and E. Paquette. Global fluctuations for linear statistics of\ beta-jacobi ensembles. Arxiv preprint arXiv:1203.6103, 2012.

[7] A. Edelman and B.D. Sutton. From random matrices to stochastic operators. Journal of Statistical Physics, 127(6):1121–1165, 2007.

[8] P.J. Forrester. Log-gases and random matrices. Number 34. Princeton Univ Pr, 2010.

[9] A. Guionnet and B. Zegarlinski. Lectures on logarithmic sobolev inequalities. Séminaire de Probabilités, XXXVI, 1801:1–134, 2002.

[10] D. Holcomb and G.R.M. Flores. Edge scaling of the\ beta-jacobi ensemble. Arxiv preprint arXiv:1203.4170, 2012.

[11] K. Johansson. Shape fluctuations and random matrices. Communications in mathematical physics, 209(2):437–476, 2000.

[12] M. Ledoux. The concentration of measure phenomenon, volume 89. Amer Mathematical Society, 2001.

[13] M. Ledoux. Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. the continuous case. Electronic Journal of Probability, 9:177–208, 2004.

[14] M. Ledoux. Deviation inequalities on largest eigenvalues. Geometric aspects of functional analysis, pages 167–219, 2007.

[15] M. Ledoux. A recursion formula for the moments of the gaussian orthogonal ensemble. In Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, volume 45, pages 754–769. Institut Henri Poincaré, 2009.

[16] M. Ledoux, B. Rider, et al. Small deviations for beta ensembles. Electronic Journal of Probability, 15:1319–1343, 2010.

[17] VA Marčenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. Mathematics of the USSR-Sbornik, 1:457, 1967.

[18] M. Noyes. Spectral Properties of the General β-Hermite and β-Laguerre Ensembles in the Limit $\beta \to \infty$. PhD thesis, University of Colorado at Boulder, 2011.

[19] F. Qi. Bounds for the ratio of two gamma functions. RGMIA Research Report Collection, 11(3), 2010.

[20] J. Ramírez, B. Rider, and B. Virág. Beta ensembles, stochastic airy spectrum, and a diffusion. Arxiv preprint math/0607331, 2006.

[21] A. Soshnikov. Universality at the edge of the spectrum in wigner random matrices. Communications in mathematical physics, 207(3):697–733, 1999.

[22] B.D. Sutton. The stochastic operator approach to random matrix theory. PhD thesis, Massachusetts Institute of Technology, 2005.

[23] C.A. Tracy and H. Widom. Level-spacing distributions and the airy kernel. Communications in Mathematical Physics, 159(1):151–174, 1994.

[24] C.A. Tracy and H. Widom. Level spacing distributions and the bessel kernel. Communications in mathematical physics, 161(2):289–309, 1994.

[25] C.A. Tracy and H. Widom. On orthogonal and symplectic matrix ensembles. Communications in Mathematical Physics, 177(3):727–754, 1996.

[26] C.A. Tracy and H. Widom. Asymptotics in asep with step initial condition. Communications in Mathematical Physics, 290(1):129–154, 2009.

[27] E.P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. The Annals of Mathematics, 62(3):548–564, 1955.