POINT PROCESS LIMITS OF RANDOM MATRICES

By

Diane Holcomb

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

Doctor of Philosophy

(Mathematics)

at the

UNIVERSITY OF WISCONSIN – MADISON

2015

Date of final oral examination: April 29, 2014

The dissertation is approved by the following members of the Final Oral Committee:

Professor Benedek Valkó, Associate Professor, Mathematics (advisor)

Professor David F. Anderson, Assistant Professor, Mathematics

Professor Timo Seppäläinen, Professor, Mathematics

Professor Philip Matchett Wood, Assistant Professor, Mathematics

Professor Saverio Spagnolie, Assistant Professor, Mathematics

Abstract

The Gaussian Unitary and Orthogonal ensembles are among the most studied models in random matrix theory. These models can be included in a one parameter family of point processes with n points indexed by β called the β-Hermite (or Gaussian) ensemble. These, and two related ensembles, the β-Laguerre and β-Jacobi, can be rescaled so that on average the number of points in a fixed interval stays the same as n grows. The resulting limit is a point process, which we will call a local limit. The work included here can be divided into two types of results. The first is proof of convergence to the local limits at the edge of the β-Jacobi ensemble. The second type of result is on properties of the limiting point processes themselves.

The β-Jacobi ensemble can exhibit two different types of behaviors at its edge. These behaviors give rise to two different limiting point processes: The $\mathrm{Bess}_{a,\beta}$ process at the hard edge, and the Airy_β at the soft edge. We prove convergence to the hard edge and soft edge under appropriate conditions. This is joint work with Gregorio Moreno-Flores and is published, see [32].

In the interior of the β-ensembles we get a different type of local limit, the Sine_β process. This process is shift invariant and the expected number of points in an interval of length λ is $\frac{\lambda}{2\pi}$. We prove a large deviation result on the probability of seeing approximately $\rho\lambda$ points as the size of the interval tends to infinity. This is joint work with Benedek Valkó and has been submitted for publication, see [33]. Lastly we show a similar large deviation result and a central limit theorem result for the number of points in the $\mathrm{Bess}_{a,\beta}$ process in the interval $[0, \lambda]$ as λ tends to infinity. This work is as yet unpublished.

Acknowledgements

First I would like to thank my committee and the entire probability group at UW-Madison for their support, encouragement, and all around decency as humans. You are one of the best groups of people a girl could hope to work with. I would like to thank my coauthor Gregorio Moreno-Flores for his work on the project that is chapter 3.

For my sanity: I would like to thank my parents for the almost infinite number of long phone conversations. Many thanks to my sister Allison for being a point at which I find my center, having you nearer then 500 miles away for the last two years has been a blessing. To Matt (Davis), Andrew, Nicos, Devon, Claire, Paul, and Emily for making me laugh when I was crying, feeding me when I wasn't up to it, dodging flying forks, being blatantly inappropriate, being my bunny, feeding me chocolate donuts and prosecco, and responding to texts demanding hugs. You all are, or have been, a part of my foundation. It gets a bit wobbly without you.

Many thanks also to the many other people who have been truly wonderful friends. You are the yardstick against which I measure myself, and the part of me of which I am most proud.

Finally, and most importantly, I would like to thank my advisor Benedek Valkó. The best decision I made in all of graduate school was taking your class my first semester here. Without that the rest of it probably wouldn't have happened at all. For putting up with my feast and famine work schedule, my occasional insanity, my managing disposition, my millions of emails (which frequently contradicted the one directly before it), and my many other idiosyncrasies, I don't have the words to thank you enough. Your guidance and support mean the world to me.

Contents

Chapter 1

Introduction and Background

The study of random matrices grew out of the work of Wishart (1920's) in statistics, and gained popularity with the work of Wigner (1950's) in nuclear physics (see e.g. [58], [57]). In the past 60 years random matrices remained objects of interest both from a purely mathematical perspective and for their deep connections and applications to other areas, both in and outside of mathematics. Mathematical connections include number theory, knot theory, and algebraic geometry. Applications of random matrices are also found in physics, statistics, finance, electrical engineering, and several other fields. The applications can be found in such varying areas as microbiology where they are used to model RNA folding, and condensed matter physics where they have been used to study the conductance of new materials (see [2]).

The early models were mostly Hermitian and symmetric random matrices and remain a focus of the field. One of the classical models, introduced by Wigner, is obtained by taking A to be a square matrix whose entries are independent standard normal random variables ($\mathcal{N}(0,1)$), and considering the matrix $M = (A + A^T)/2$. The matrix M is a sample from the Gaussian Orthogonal Ensemble (GOE). Related models can be constructed with complex or quaternion entries and using the appropriate adjoint, the resulting matrix models are called the Gaussian Unitary and Gaussian Symplectic Ensembles (GUE, GSE). These models remain central to the study of random matrices because they are some of the most approachable models to work with and share characteristics with a broad range of other random matrix models. Other commonly studied random matrix models include covariance (also called Wishart) matrices, MANOVA matrices, and unitary matrices chosen randomly according to a Haar measure.

In the remainder of the introduction we will use the Gaussian ensembles to introduce some of the types of problems of interest in random matrices. We will then use them to introduce a generalization of random matrices known as β-ensembles. Lastly we will give a brief overview of the work covered in this thesis.

1.1 Limits of the Gaussian ensembles

The type of question of interest here is what happens to the spectrum of the matrices as the size of the matrix tends to infinity. There are two different scales on which we consider this problem. The first of these is to look at the behavior of the eigenvalues collectively, where and how densely they accumulate. The second is to look at the local interaction of the eigenvalues.

One way of describing the collective behavior of the eigenvalues is to look at the empirical spectral measure of the matrix. This is the measure that places a weight of $1/n$ on each eigenvalue. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalues of a random matrix, then we define the *empirical spectral measure* to be

$$\nu_n(x) = \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k}(x).$$

Now specialize to the Gaussian ensembles, that is let the λ_i be the eigenvalues of an n by n GOE, GUE, or GSE matrix. Then as n tends to infinity we can check that we expect the largest and smallest eigenvalues to be around $\pm 2\sqrt{n}$. If we rescale the spectrum by $1/\sqrt{n}$ then we get the following theorem:

Theorem 1.1 (Wigner)**.** *Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigenvalue of a GOE, GUE, or GSE matrix, and let*

$$\mu_n = \frac{1}{n}\sum_{k=1}^{n} \delta_{\lambda_k/\sqrt{n}}(x)$$

then

$$\mu_n \Rightarrow \rho_{sc} \quad a.s. \qquad\qquad \textit{where} \qquad\qquad \frac{d\rho_{sc}}{dx} = \frac{1}{2\pi}\sqrt{4-x^2}.$$

The meaning of this statement can be seen from Figure 1.1, which shows a histogram of the eigenvalues of a 1000 by 1000 GUE matrix scaled by $1/\sqrt{n}$. From this picture we can clearly see the emergence of the semicircle mentioned.

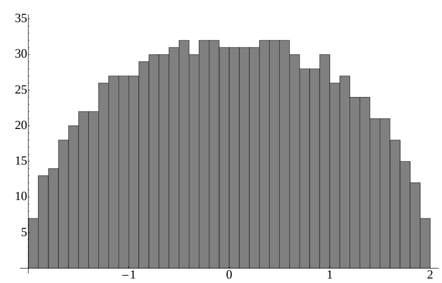

Figure 1.1: Histogram of rescaled eigenvalues of a 1000×1000 GUE matrix

Remark 1.2. *The Wigner semicircle law holds more generally then stated above. Take any $n \times n$ symmetric or Hermitian matrix with entries below the diagonal i.i.d. with mean 0 and variance 1, and diagonal entries i.i.d. mean 0 and variance 1. Then when rescaled by $1/\sqrt{n}$ the empirical spectral density converges to the semicircle distribution. (See e.g. [3])*

This type of limit describes the global behavior of the eigenvalues, but tells you nothing about the interactions between the eigenvalues. In order to look at what happens on a local scale we need to rescale so the expected number of eigenvalues in an interval remains constant. In the case of the Gaussian ensembles we have n eigenvalues on an interval of size $4\sqrt{n}$, this suggests that we need to look at a subinterval on the order of $1/\sqrt{n}$. Indeed, in the limit, after rescaling the eigenvalues up by $\sqrt{n}$ we arrive at a point process. In the case of $\beta = 2$, if we take Λ_n to be the set of eigenvalues, then the precise statement is that $\frac{1}{\pi}\sqrt{n}\Lambda_n$ converges in distribution to the Sine process, which may be described through its n-point correlation functions. These are given by determinantal formulas with entries $p(x_i, x_j) = \frac{\sin \pi(x_i - x_j)}{\pi(x_i - x_j)}$. (see e.g. [28], and for more on determinantal point processes see [34]).

Observe that by looking at $\frac{1}{\pi}\sqrt{n}\Lambda_n$ we are looking at the local behavior of the eigenvalues near 0, and notice that 0 is in the interior of the spectrum (that is, in the middle of the semicircle formed by the eigenvalues). We refer to this as bulk of the spectrum, and the Sine process as the bulk limiting process. If we rescale appropriately at another point in the interior of the spectrum we find this same limiting process. In other words the local behavior is the same everywhere in the interior. In the case of the GUE this would be anywhere in $(-2\sqrt{n}, 2\sqrt{n})$ as we can see from Figure 1.1.

We can find a similar result at the edge of the spectrum, though with a different limiting process. Notice that at the very edge of the spectrum from Wigner's semicircle law we get that the number of points in the interval $[-2\sqrt{n}, (-2+\epsilon)\sqrt{n}]$ is approximately $n\int_{-2}^{-2+\epsilon} \rho_{sc}(x)dx \sim nC\epsilon^{3/2}$. This suggests that in order to look at the local behavior we need ϵ on the order of $n^{-2/3}$, which means we need to look at the interval $[-2\sqrt{n}, -2\sqrt{n} + n^{-1/6}]$. Indeed, appropriately centered and rescaled by $n^{1/6}$ the bottom eigenvalues converge the Airy_2 process, which can again be described through its correlation functions. This time the determinantal formulas are written in terms of Airy functions. See [28] for further details.

1.2 β-ensembles

The Gaussian ensembles introduced above remain important to the study of random matrices in part because we can write down an explicit joint density for the eigenvalues. Let M_n be an $n \times n$ GOE, GUE, or GSE matrix, then the eigenvalue have joint density

$$p_H(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_{n,\beta}} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^{n} e^{-\frac{\beta}{4}\lambda_i^2} \tag{1.1}$$

where β is equal to 1, 2 and 4 for the GOE, GUE, and GSE respectively. Note, that the constant $\beta/4$ in the exponential can be easily changed via linear scaling, but not the exponent in the first product. As a generalization of the Gaussian ensembles we can consider the density in (1.1) for any $\beta > 0$, this is called the Gaussian (or Hermite) β-ensemble. We will denote this by $H(n, \beta)$.

The finite point process $H(n, \beta)$ is no longer directly related to the original matrix models, but there does exist a related tridiagonal matrix model whose eigenvalues have the appropriate distribution. The model originally introduced by Dumitriu and Edelman [17] is as follows: Let

$$H_\beta \sim \frac{1}{\sqrt{2}} \begin{bmatrix} N(0,2) & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & N(0,2) & \chi_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & N(0,2) & \chi_\beta \\ & & & \chi_\beta & N(0,2) \end{bmatrix} \tag{1.2}$$

then the joint density of the eigenvalues of H_β is given by (1.1). The tridiagonal matrix model in integral to the study of the β-ensemble. There is another interpretation of this model in terms of log gases. In particular the eigenvalue density given in (1.1) can be identified with the Boltzmann factor of a log-gas model, more work on this can be found in the physics literature [28].

The β-Hermite ensemble is one example of a family of distributions that may be considered. More generally if we have a joint density function of the form

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,\beta}} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^\beta \prod_{i=1}^{n} e^{V(\lambda_i)}$$

for some function V, we call a set of points distributed according to such a function a β-ensemble. There are several other classical matrix models that also generalize to β-ensembles. These include, but are not limited to Wishart matrices with Gaussian entries which generalize to the β-Laguerre ensemble, and MANOVA matrices which generalize to the β-Jacobi ensemble.

The model considered by Wishart [58] can be described as follows: Let $\mathcal{M}_{n\times m}$ denote the space of $n \times m$ matrices with complex (resp. real) valued entries. Take a matrix $X \in \mathcal{M}_{n\times m}$ with independent complex (resp. real) Gaussian entries; then $M = XX^T \in \mathcal{M}_{n\times n}$ is said to have a (n, m)-Wishart distribution. It corresponds to the sample covariance matrix for a sample drawn from a multivariate complex (resp. real) normal distribution. The joint eigenvalue density given by

$$p_L(\lambda_1, \lambda_2, \dots, \lambda_n) = \frac{1}{Z_{\beta,n,m}} \prod_{i=1}^{n} \lambda_i^{\frac{\beta}{2}(m-n+1)-1} e^{-\frac{\beta}{2}\lambda_i} \prod_{j<k} |\lambda_j - \lambda_k|^\beta \tag{1.3}$$

where $\beta = 1, 2$ for real and complex initial entries respectively. Taking any $\beta > 0$ we call the resulting object the β-Laguerre ensemble, which we will denote by $L(n, m, \beta)$. As in the β-Hermite case, there is a tridiagonal matrix model whose eigenvalues are distributed according to (1.3), see [17] for more details.

The MANOVA matrix model also comes from statistics. We take an n_1 by n matrix M with each entry independently drawn from the standard normal distribution, and consider the matrix $X = M^T M$. This is one of the n by n Hermitian matrix models introduced by Wishart. If we assume $n_1 \geq n$ we know the resulting matrix is almost surely invertible and so $A = X^{1/2}(X+Y)^{-1}X^{1/2}$ where Y is constructed in the same manner with parameters n_2, n is well defined. Consideration of this type of matrix first arose in statistics (MANOVA, or multivariate analysis of variance, uses the base matrix model to study the interdependence of several dependent and independent variables). The resulting matrix A is again Hermitian and joint density function of its eigenvalues is given by

$$p_J(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_{\beta,n,n_1,n_2}} \prod_{i=1}^{n} \lambda_i^a (1-\lambda_i)^b \prod_{j<k} |\lambda_j - \lambda_k|^\beta \, \mathbf{1}_{\{\lambda_i \in [0,1], \forall i\}} \tag{1.4}$$

where $a = \frac{\beta}{2}(n_1 - n + 1) - 1, b = \frac{\beta}{2}(n_2 - n + 1) - 1$, Z is a constant dependent on the parameters, $n_1, n_2 \geq n$ and $\beta = 1$. If instead of drawing from the standard normal distribution we had chosen a complex normal distribution, we would have had the same joint density function with $\beta = 2$. These cases were extensively studied by Johnstone in [37].

Note that the eigenvalues lie inside the interval $[0, 1]$, unlike the Gaussian or Wishart cases where the spectrum is unbounded. Taking $\beta > 0$ the resulting object is called the β-Jacobi ensemble, we will denote this ensemble by $J(n, n_1, n_2, \beta)$. The corresponding tridiagonal matrix model is introduced in Section 3.1.1.

1.3 Limits of β-ensembles

For the β-ensembles we can study the same type of questions as in the case of the GUE. As with the GUE we will begin by looking at the limits of the empirical spectral density.

We will then take a look the local limit that shows up in the bulk of the β-ensembles. Lastly we will look at the point process limits at the edge of the ensembles.

1.3.1 Limits of the empirical spectral measures

Recall that in the case of the Gaussian ensembles (the GOE/GUE/GSE models) we had that appropriately rescaled the empirical spectral density converged in distribution to the semicircle law (see Theorem 1.1). There are similar theorems that hold for the Wishart and MANOVA matrices, but with different limiting distributions. The empirical spectral densities of the β-ensembles converge to the same distributions you get in the original matrix models. More precisely we have the following results:

Theorem 1.3 (see e.g. [28]). *1. Let $\lambda_1, \lambda_2, ..., \lambda_n \sim H(n, \beta)$, then*

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i/\sqrt{n}}(x) \Rightarrow \rho_{sc}(x) \ a.s. \ as\ n \to \infty. \tag{1.5}$$

2. Let $\lambda_1, \lambda_2, ..., \lambda_n \sim L(n, m, \beta)$ and assume $n/m \to \gamma \in (0, 1]$, then

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i/m} \Rightarrow \rho_{mp}(x) \ a.s. \ as\ n \to \infty \tag{1.6}$$

where ρ_{mp} is the Marchenko-Pastur distribution with

$$\frac{d\rho_{mp}}{dx} = \frac{1}{2\pi\gamma x}\sqrt{(b-x)(x-a)}\mathbf{1}(a \le x \le b), \quad a = (1-\sqrt{\gamma})^2, \quad b = (1+\sqrt{\gamma})^2. \tag{1.7}$$

3. Let $\lambda_1, \lambda_2, ..., \lambda_n \sim J(n, n_1, n_2, \beta)$ with $n_1, n_2 > n$, $n_1/n \to \gamma_1 \in [1, \infty)$, and $n_2/n \to \gamma_2 \in [1, \infty)$, then $\frac{1}{n}\sum_i \delta_{\lambda_i}$ converges weakly to $\rho(x)$ a.s. as $n \to \infty$, where

$$\rho(x) = \frac{2\pi}{\gamma_1 + \gamma_2} \cdot \frac{\sqrt{(\Lambda_+ - x)(x - \Lambda_-)}}{x(1-x)}.$$

Here $\Lambda_\pm$ denotes the upper and lower edges of the spectrum which are given by

$$\Lambda_\pm = \left(\frac{\sqrt{\gamma_1(\gamma_1+\gamma_2-1)}}{\gamma_1+\gamma_2} \pm \frac{\sqrt{\gamma_2}}{\gamma_1+\gamma_2} \right)^2 . \tag{1.8}$$

These limiting densities help us identify the interior (or bulk) and edges of the spectrum in the limit. They also help to find the scale of which the local interactions will be visible. We use this in the discussion of the local limits the follows.

1.3.2 The local limits in the bulk

Recall that by local limit we are referring to a point process that captures the local interactions of the eigenvalues. In the case of the GUE we had that rescaled by $\sqrt{n}$ in the limit the eigenvalues converged in distribution to a determinantal point process we called the Sine process. There is a similar statement for the β-Hermite ensemble. In this case the limiting point process is dependent on β and is denoted by Sine_β. This process will be characterized through its counting function rather then its correlation functions. Characterization of this process and proof of convergence in the case of the β-Hermite ensemble was done by Valkó and Virág [54].

The Sine_β process as the bulk limit is not unique to the β-Hermite ensemble. Appropriately rescaled this same limit if found in the bulk of the β-Laguerre and -Jacobi ensembles ([35], [7]). Notice that the only dependence in the limit is on β, meaning that, in the limit, the extra parameters in these ensembles have no impact on the local interactions in interior. The behavior at the edge of the spectrum is more diverse.

1.3.3 Limits at the edge

Recall that at the edge of the GUE we had that centered and rescaled by $n^{1/6}$ the eigenvalues converged in distribution to the Airy$_2$ process. As with the bulk we have a similar statement for the edges of the β-Hermite ensemble. In particular, if Λ_n denotes the collection of the ordered eigenvalues of $H(n, \beta)$, then $n^{1/6}(2\sqrt{n} - \Lambda_n)$ converges in distribution to a point process Airy$_\beta$, which depends only on β. This limiting point process can be characterized in two ways. The first is that it is the ordered eigenvalues of a stochastic differential operator. The second is again through its counting function. This type of behavior will be called a soft edge [49].

This limiting point process, unlike the limit in the bulk, is not the only process found at the edge of the β-Laguerre and -Jacobi ensembles. It is also possible that these ensembles display a different type of behavior at the edge of the spectrum. To see why this might be expected consider the Marchenko-Pastur distribution which is found as a limit of the empirical spectral density of the β-Laguerre ensemble (1.7). This limiting distribution can exhibit two very different types of behavior at the lower edge of the support. In the case where $n/m \to 1$ we get a limiting density with

$$\frac{d\rho_{mp}}{dx} = \frac{1}{2\pi\gamma\sqrt{x}}\sqrt{2-x}\mathbf{1}(0 \leq x \leq 2).$$

This gives us an asymptote at $x = 0$. Returning to the distribution of the β-Laguerre ensemble, this means that the lower edge is converging to 0, but the joint density by construction requires that the eigenvalues be nonnegative, which imposes a hard constraint at the origin. Indeed in the case where $m = n + a$ when appropriately rescaled the smallest k eigenvalues of the β-Laguerre ensemble converge in distribution to a point process dependent on β and a. We will call this process the 'Stochastic Bessel Process' and denote it by $\mathrm{Bess}_{a,\beta}$. This names comes from the fact that for $\beta = 2$ this process can again be characterized by its n-point correlation functions, this time the necessary determinant being composed of Bessel functions (see e.g. [28]). Characterization of the limiting point process as well as the proof of convergence was done by Ramírez and Rider in [48].

For $m/n \to \gamma \neq 1$ the lower edge of the Marchenko-Pastur distribution behaves like the $\sqrt{x}$ function near 0. This is the same as the behavior at the top and bottom of the semicircle in the β-Hermite limit. In this case, that is $m/n \to \gamma \neq 1$, the appropriately rescaled lower edge of the β-Laguerre ensemble converges to the Airy_β process [49].

The regime in between these two, that is $m/n \to 1$, but $m = n + a_n$ where $a_n \to \infty$, should fall into the case of the soft edge, but the actual result remains unproven. The upper edge of the β-Laguerre ensemble falls into the soft edge regime regardless of the limit of m/n.

The last thing that remains is to discuss the limiting behavior at the edge of the β-Jacobi ensemble. As with the β-Laguerre ensemble, this ensemble can exhibit both soft and hard edge behavior, but the behavior can change at both its upper and lower edges. The n_1 parameter determines the behavior at the lower edge, while the n_2 parameter determines the behavior at the upper edge. In particular, if $n_1 = n + a$ then the lower edge converges to $\mathrm{Bess}_{a,\beta}$, and similarly for the upper edge if $n_2 = n + a$. If $n_1/n \to \gamma_1 \neq 1$, then the lower edge converges to Airy_β, and similarly for the upper edge if $n_2/n \to \gamma_2 \neq 1$. As in the β-Laguerre ensemble, the transitional case remains unproven. The proof of the limits in the β-Jacobi case is one part of the work in this thesis.

1.3.4 Universality of the point process limits

The universality conjecture is that these local limits are the correct limits for any reasonable matrix model, and this does not depend on the on the limiting empirical spectral density. For example if one takes M with $M = M^*$ with $m_{i,j}$ any complex random variables with mean 0 and variance 1 for $i < j$, and $m_{i,i}$ real with mean 0 and variance 1, then for Λ_n the eigenvalues of M we have $2\sqrt{n}\Lambda_n \Rightarrow \mathrm{Sine}_2$. This has been proved under certain conditions by Erdős, Yau, and several coauthors, and independently by Tao and Vu, see [25, 24, 23, 52]. Related work shows that the bulk limit also holds for a wide family of β-ensembles [7]. It has also been proved that the edge limit holds for a large family of β-ensembles [8, 42]. In this sense we can understand these point process limits which describe the local interaction of the eigenvalues to be a type of universal behavior.

1.4 Summary of Results

The results of this thesis can be divided into two types. The first of these is the proof of convergence of the β-Jacobi ensembles to the local limits at the edge of the spectrum. The second of these are on properties of the limiting point processes. This includes large deviation results for the Sine_β and $\mathrm{Bess}_{a,\beta}$ process, as well as a CLT for the $\mathrm{Bess}_{a,\beta}$ process. In particular we show the following:

For the β-Jacobi ensemble, with joint density given by (1.4) we show that when appropriate centered and rescaled we get convergence to the limiting point processes for both the hard and soft edge. For the soft edge this means that for $\liminf n_1/n > 1$ the ordered eigenvalue when appropriated centered and rescaled converge in distribution to the stochastic Airy process. The result will be stated precisely for the upper edge of the spectrum, that is for $\liminf n_2/n > 1$, but symmetry gives the lower edge as well. For the hard edge we have that for $n_1 - n \to a$, the ordered eigenvalue of the β-Jacobi spectrum converge to the stochastic Bessel process with parameter a. These results have been published previously in [32].

The large deviation results are on the counting functions of the Sine_β and stochastic Bessel processes. The results state that the normalized counting function $N(\lambda)/\lambda$ of the Sine_β process satisfies a large deviation principle with scale λ^2 and a good rate function $\beta I_{\text{Sine}}(\rho)$. That is $P(N(\lambda)/\lambda \sim \rho) \sim e^{-\lambda^2 \beta I_{Sine}(\rho)}$. This work has been submitted for publication and can be found in [33].

The large deviation result for the $\text{Bess}_{a,\beta}$ process is similar. The counting function of the stochastic Bessel process on the interval $[0, \lambda]$ normalized by $\sqrt{\lambda}$ satisfies a large deviation principle with scale λ and good rate function $\beta I_{\text{Bess}_{a,\beta}}$. The central limit type result is for the number of points in the interval $[0, \lambda]$ as $\lambda \to \infty$. This work has not been published yet.

These results will be stated precisely in the next chapter.

Chapter 2

Results

2.1 Point Process Limits of the β-Jacobi Ensemble

2.1.1 Soft edge limit

We begin by defining the "stochastic Airy operator" (SAO$_\beta$) . Let

$$\mathcal{H}_\beta = -\frac{d^2}{dx^2} + x + \frac{2}{\sqrt{\beta}} b'(x) \tag{2.1}$$

where we take b' to be a white noise. A precise definition and many properties of this operator can be found in [49]. We review the necessary ones below.

For our purposes it is sufficient to define an eigenfunction/eigenvalue pair in the following way: Let

$$L^* = \left\{ f \in L^2[0,\infty) | f(0) = 0, f' \text{ exists a.e. and } \int_0^\infty (f')^2 + (1+x) f^2 dx < \infty \right\},$$

then (φ, λ) is an eigenvalue/eigenfunction pair for $\mathcal{H}_\beta$ if $\|\varphi\|_2 = 1, \varphi \in L^*$ and

$$\varphi''(x) = \frac{2}{\sqrt{\beta}} \varphi(x) b'(x) + (x - \lambda)\varphi(x) \tag{2.2}$$

holds in the sense of distributions. This may be written as

$$\varphi'(x) - \varphi'(0) = \frac{2}{\sqrt{\beta}} \varphi(x) b(x) - \frac{2}{\sqrt{\beta}} \int_0^x \varphi'(t) b(t) dt + \int_0^x (t - \lambda)\varphi(t) dt. \tag{2.3}$$

In this sense, the set of eigenvalues is a deterministic function of the Brownian path b. Moreover the eigenvalues are "nice" in the following sense:

Theorem 2.1. *[49] With probability one, the eigenvalues of $\mathcal{H}_\beta$ are distinct (of multiplicity 1) with no accumulation point, and for each $k \geq 0$ the set of eigenvalues of $\mathcal{H}_\beta$ has a well defined $(k+1)$st lowest element $\Lambda_k(\beta)$.*

The Airy$_\beta$ point process is given by the eigenvalues of $\mathcal{H}_\beta$. Our first result shows that the spectrum of the Jacobi ensemble near to the soft edge converges to this point process after appropriate scaling and centering.

Let us introduce some notation: Take

$$c^2 = \frac{n_1}{n_1+n_2}, \qquad s^2 = \frac{n_2}{n_1+n_2} \tag{2.4}$$

$$\tilde{c}^2 = \frac{n}{n_1+n_2}, \qquad \tilde{s}^2 = \frac{n_1+n_2-n}{n_1+n_2} \tag{2.5}$$

with $c, s, \tilde{c}$ and $\tilde{s}$ all nonnegative. Under this notation we have that the expected edge of the spectrum is given by

$$\Lambda_\pm = (c\tilde{s} \pm s\tilde{c})^2.$$

We define our scaling factor to be

$$\alpha_n = \frac{m_n^2}{cs\tilde{c}\tilde{s}}, \qquad \text{where} \qquad m_n = \left[\frac{cs\tilde{c}\tilde{s}\sqrt{n_1+n_2}}{\tilde{c}\tilde{s}(c^2-s^2)+cs(\tilde{c}^2-\tilde{s}^2)}\right]^{2/3}. \tag{2.6}$$

We now state the scaling limit near the soft edge.

Theorem 2.2. *Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ denote the ordered eigenvalues of $J(n, n_1, n_2, \beta)$, and assume $\liminf n_2/n > 1$; then*

$$\Big(\alpha_n(\Lambda_+ - \lambda_\ell)\Big)_{\ell=1,\ldots,k} \Rightarrow \Big(\Lambda_0(\beta), \ldots, \Lambda_{k-1}(\beta)\Big)$$

jointly in law for any fixed $k < \infty$, as $n \to \infty$.

Remark 2.3. *This theorem describes the limiting behavior of the β-Jacobi ensemble in the upper soft edge situation. This is sufficient to determine the behavior in the lower soft edge because $J(n, n_1, n_2, \beta)$ is symmetric in n_1 and n_2. That is, the reflection of the density of $J(n, n_1, n_2, \beta)$ with respect to $x = 1/2$ is the corresponding density for $J(n, n_2, n_1, \beta)$.*

Remark 2.4. *In the situation where n_1 and n_2 are constant multiples of n, then $\alpha_n = cn^{2/3}$ for some constant c. One can compare this with the scaling exponents in the Tracy-Widom result. In fact in the case where $\liminf n_2/n \geq 1 + \epsilon$, then with a little bit of work one can show that there exist constants c_1 and c_2 depending only on ϵ so that $c_1 n^{1/3} \leq m_n \leq c_2 n^{1/3}$.*

Remark 2.5. *While we expect a soft edge type result for β-Jacobi spectrum in the case where $n_2 = n + a_n$ and $a_n \to \infty$ we restrict to the case where $\liminf n_2/n > 1$. We make this restriction on n_2 because, as the previous remark suggests, the order of m_n can change substantially which will render many computations invalid. A similar gap exists in the β-Laguerre case for the lower soft edge.*

2.1.2 Hard edge limit

We recall the definition of the stochastic Bessel operator, studied by Ramírez and Rider in [48] to describe the limit of the Laguerre ensemble in the hard edge case. The operator acts on functions on $\mathbb{R}_+$ and is given by:

$$\mathfrak{G}_{\beta,a} = -\exp\left[(a+1)x + \frac{2}{\sqrt{\beta}}b(x)\right] \cdot \frac{d}{dx}\left(\exp\left[-ax - \frac{2}{\sqrt{\beta}}b(x)\right]\frac{d}{dx}\right),$$

with Dirichlet boundary conditions at 0 and Neumann conditions at infinity, where $b(x)$ is a Brownian motion, $a > -1$ and $\beta > 0$. This can be formally rewritten as

$$-\mathfrak{G}_{\beta,a} = e^x \left(\frac{d^2}{dx^2} - (a + \frac{2}{\sqrt{\beta}} b'(x)) \frac{d}{dx} \right),$$

where $b'(x)$ is white noise. The inverse operator as given by Ramírez and Rider [48] is:

$$(\mathfrak{G}_{\beta,a}^{-1}\psi)(x) \equiv \int_0^\infty \left(\int_0^{x\wedge y} e^{az + \frac{2}{\sqrt{\beta}}b(z)} dz \right) \psi(y) e^{-(a+1)y - \frac{2}{\sqrt{\beta}}b(y)} dy. \tag{2.7}$$

The operator $\mathfrak{G}_{\beta,a}^{-1}$ is non-negative symmetric in $L^2[\mathbb{R}_+, m]$ where

$$m(dx) = e^{-(a+1)x - \frac{2}{\sqrt{\beta}}b(x)} dx.$$

We may study the eigenvalues of $\mathfrak{G}_{\beta,a}$ through the study of the eigenvalues of the integral operator by using the fact that if λ is an eigenvalue of $\mathfrak{G}_{\beta,a}$ its reciprocal is an eigenvalue of the inverse. Moreover, it can be shown that the spectrum defines a simple point process as desired.

Theorem 2.6. *[48] With probability one, when restricted to the positive half-line with Dirichlet boundary condition (at the origin), $\mathfrak{G}_{\beta,a}$ has a discrete spectrum of simple eigenvalues $0 < \Lambda_0(\beta, a) < \Lambda_1(\beta, a) < \cdots \uparrow \infty$.*

We will show that the Bessel operator will describe the limiting spectrum of the (β, n, n_1, n_2)-Jacobi ensemble in the hard edge case. For convenience and in analogy to the β-Laguerre ensemble, we will focus on the lower hard edge. We can see from Remark 2.3 that this is indeed sufficient to determine the upper edge as well.

Theorem 2.7. *Let $0 < \lambda_0 < \lambda_1 < \cdots < \lambda_{n-1}$ be the ordered eigenvalues of $J(n, n_1, n_2, \beta)$, $m_n = nn_2$ and $n_2 > n$. Assume that $(n_1 - n) \to a \in (-1, \infty)$, then*

$$\Big(m_n\lambda_0, m_n\lambda_1, ..., m_n\lambda_k \Big) \Rightarrow \Big(\Lambda_0(\beta, a), \Lambda_1(\beta, a), ..., \Lambda_k(\beta, a) \Big)$$

jointly in law, for any fixed $k < \infty$ as $n \to \infty$.

Remark 2.8. *Previous work on both the hard and the soft edge of the β-Jacobi was done through a coupling with the β-Laguerre ensemble by Jiang [36]. This work covers the case where $n/n_1 \to \gamma \in (0,1]$ and $n = o(\sqrt{n_2})$. Our work extends the results to all cases for which* $\liminf n_2/n > 1$.

2.2 Properties of Point Process Limits

2.2.1 Large deviations for the Sine_β process

The Sine_β process can be described through its counting function using a system of stochastic differential equations. Consider the system

$$d\alpha_\lambda = \lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} dt + \mathrm{Re}\left[(e^{-i\alpha_\lambda} - 1)(dB_1 + idB_2)\right], \qquad \alpha_\lambda(0) = 0, \quad t \in [0, \infty) \tag{2.8}$$

where B_1, B_2 are independent standard Brownian motions. Note, that this is a one-parameter family of SDEs driven by the same complex Brownian motion. In [54] it was shown that $N_\beta(\lambda) = \lim_{t\to\infty} \frac{1}{2\pi}\alpha_\lambda(t)$ exists almost surely and it is an integer valued monotone increasing function in λ. Moreover, the function $\lambda \to N_\beta(\lambda)$ has the same distribution as the counting function of the Sine_β process, i.e. the distribution of the number of points in $[0, \lambda]$ for $\lambda > 0$ is given by that of $N_\beta(\lambda)$.

Note, that for any fixed λ the process α_λ satisfies the SDE

$$d\alpha_\lambda = \lambda \frac{\beta}{4} e^{-\frac{\beta}{4}t} dt + 2\sin(\alpha_\lambda/2) dB_t, \qquad \alpha_\lambda(0) = 0, \quad t \in [0, \infty) \tag{2.9}$$

where $B_t = B_t^{(\lambda)} = \int_0^t \mathrm{Re}\big[-ie^{-\frac{1}{2}i\alpha_\lambda(s)} d(B_1 + idB_2)\big]$ is a standard Brownian motion which depends on λ. Thus, if we are interested in the number of points in a given interval $[0, \lambda]$ then it is enough to study the SDE (2.9) instead of the system (2.8).

Using the SDE characterization of the Sine_β process one can show that it is translation invariant with density $(2\pi)^{-1}$ (see [54]). In particular, in a large interval $[0,\lambda]$ one expects roughly $(2\pi)^{-1}\lambda$ points. In [43] the authors refined this by showing that $N_\beta(\lambda)$ satisfies a central limit theorem, it is asymptotically normal with mean $\frac{\lambda}{2\pi}$ and variance $\frac{2}{\beta\pi^2}\log\lambda$.

The goal here is to characterize the large deviation behavior of $N_\beta(\lambda)$. We find the asymptotic probability of seeing an average density different from $(2\pi)^{-1}$ on a large interval. We show that $\lambda^{-1}N_\beta(\lambda)$ satisfies a large deviation principle with a good rate function.

Before stating the exact form of the theorem we need to introduce a couple of notations. We will use

$$K(a)=\int_0^{\pi/2}\frac{dx}{\sqrt{1-a\sin^2 x}},\qquad E(a)=\int_0^{\pi/2}\sqrt{1-a\sin^2 x}\,dx, \tag{2.10}$$

for the complete elliptic integrals of the first and second kind, respectively. Note that there are several conventions denoting these functions, we use the one in [1]. We also introduce the following function for $a<1$:

$$\mathcal{H}(a)=(1-a)K(a)-E(a). \tag{2.11}$$

Now we are ready to state our main theorem.

Theorem 2.9. *Fix $\beta>0$. The sequence of random variables $\frac{1}{\lambda}N_\beta(\lambda)$ satisfies a large deviation principle with scale λ^2 and good rate function $\beta I_{Sine}(\rho)$ with*

$$I_{Sine}(\rho)=\frac{1}{8}\left[\frac{\nu}{8}+\rho\mathcal{H}(\nu)\right],\qquad \nu=\gamma^{(-1)}(\rho), \tag{2.12}$$

where $\gamma^{(-1)}$ denotes the inverse of the continuous, strictly decreasing function given by

$$\gamma(\nu) = \begin{cases} \frac{\mathcal{H}(\nu)}{8} \int\limits_{-\infty}^{\nu} \mathcal{H}^{-2}(x)dx, & \text{if } \nu < 0, \\ \\ \frac{1}{2\pi}, & \text{if } \nu = 0, \\ \\ \frac{\mathcal{H}(\nu)}{8} \int\limits_{1}^{\nu} \mathcal{H}^{-2}(x)dx, & \text{if } 0 < \nu < 1, \\ \\ 0, & \text{if } \nu = 1. \end{cases} \tag{2.13}$$

Roughly speaking, this means that the probability of seeing close to $\rho\lambda$ points in $[0, \lambda]$ for a large λ is asymptotically $e^{-\lambda^2 \beta I_{\text{Sine}}(\rho)}$. The precise statement is that if G is an open, and F is a closed subset of $[0, \infty)$ then

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda^2} P(\tfrac{1}{\lambda} N_\beta(\lambda) \in G) \geq -\inf_{x\in G} \beta I_{\text{Sine}}(x), \quad \limsup_{\lambda\to\infty} \frac{1}{\lambda^2} P(\tfrac{1}{\lambda} N_\beta(\lambda) \in F) \leq -\inf_{x\in F} \beta I_{\text{Sine}}(x).$$

The function γ may also be defined as the solution to the equation $4x(1-x)\gamma''(x) = \gamma(x)$ on the intervals $(-\infty, 0]$ and $[0, 1]$ with boundary conditions $\lim_{x\to 0^\pm} \gamma(x) = \frac{1}{2\pi}$, $\gamma(1) = 0$ and $\lim_{x\to-\infty} \frac{\gamma(x)}{\sqrt{|x|}} = \frac{1}{4}$. The rate function $I_{\text{Sine}}(\rho)$ is strictly convex and non-negative with $I_{\text{Sine}}(\frac{1}{2\pi}) = 0$ and $I_{\text{Sine}}(0) = \frac{1}{64}$. The function $I_{\text{Sine}}(\frac{1}{2\pi} + x)$ behaves like $-\frac{\pi^2 x^2}{4 \log(1/|x|)}$ for small $|x|$, and $I_{\text{Sine}}(\rho)$ grows like $\frac{1}{2}\rho^2 \log \rho$ as $\rho \to \infty$. These statements will be proved in Proposition B.5.

We note that the behavior of $I_{\text{Sine}}(\rho)$ near $\rho = \frac{1}{2\pi}$ is formally consistent with the already mentioned central limit theorem of $N_\beta(\lambda)$. For $\rho = \frac{1}{2\pi} + x$ with a small, but fixed $|x|$ the probability of seeing close to $\frac{1}{2\pi}\lambda + x\lambda$ points in $[0, \lambda]$ is approximately $\exp\left(-\frac{\beta\pi^2\lambda^2 x^2}{4\log(1/|x|)}\right)$. Now let us assume, that this is true even if x decays with λ, even though this regime is not covered in our theorem. If we substitute $\lambda x = \sqrt{\frac{2}{\beta\pi^2} \log \lambda} \cdot y$

(with a fixed y), then this probability would asymptotically equal to $e^{-y^2/2}$. This is in agreement with the fact that $N_\beta(\lambda)$ is asymptotically normal with mean $\frac{1}{2\pi}\lambda$ and variance $\frac{2}{\beta\pi^2}\log\lambda$.

Before moving on, a couple of historical notes are in order. In [54] the authors also show another large deviation statement for the Sine_β process regarding large intervals, namely that the asymptotic probability of not seeing any points in $[0,\lambda]$ is approximately $e^{-\frac{\beta}{64}\lambda^2}$. In [55] this result was sharpened by providing the more precise asymptotics of

$$P(N_\beta(\lambda)=0)=(\kappa_\beta+o(1))\lambda^{\upsilon_\beta}\exp\left\{-\frac{\beta}{64}\lambda^2+\left(\frac{\beta}{8}-\frac{1}{4}\right)\lambda\right\}, \qquad \text{as } \lambda\to\infty \tag{2.14}$$

with $\upsilon_\beta=\frac{1}{4}\left(\frac{\beta}{2}-\frac{2}{\beta}-3\right)$ and a positive constant κ_β whose value was not determined. Similar results have been proven before for the classical cases $\beta=1,2,4$, see e.g. [4], [53], [56], [13]. Moreover, the value of κ_β and higher order asymptotics were also established for these specific cases by [41], [22], [12]. Further extension in the classical cases include the exact asymptotics of $P(N_\beta(\lambda)=n)$ for fixed n and also for $n=o(\lambda)$. (See [53] and [28] for details.) In all of these results the main term of the asymptotic probability is $e^{-\frac{\beta}{64}\lambda^2}$. This is consistent with our result, as Theorem 2.9 and $I_{\mathrm{Sine}}(0)=\frac{1}{64}$ implies

$$\lim_{\varepsilon\to 0}\lim_{\lambda\to\infty}\frac{1}{\lambda^2}\log P(N_\beta(\lambda)\le\varepsilon\lambda)=-\frac{\beta}{64}.$$

The large deviation rate function (2.15) has been predicted using non-rigorous scaling and log-gas arguments in [20] and [27]. (See Section 14.6 of [28] for an overview.) Using the same techniques [29] treats the corresponding problem for the soft edge and hard edge limit processes of β-ensembles.

One can also study the large deviation behavior of the empirical distribution of the β-ensembles on a macroscopic level. It is known that after scaling with $\sqrt{n}$ the empirical measure of the distribution (1.1) converges to the Wigner semicircle law. In [5] the authors prove a large deviation principle for the scaled empirical measure, this describes the asymptotic probability of seeing a different density profile than the semicircle. One could consider our theorem a microscopic analogue of that result.

2.2.2 Large deviations and a CLT for the $\mathrm{Bess}_{a,\beta}$ process

The results for the $\mathrm{Bess}_{a,\beta}$ process will be stated in terms of the function $M_{a,\beta}(\lambda)$ which we define to be the number of points of the $\mathrm{Bess}_{a,\beta}$ process in the interval $[0, \lambda]$. We show both a central limit theorem and a large deviation result for this function as $\lambda \to \infty$.

Theorem 2.10. *Fix $\beta > 0, a > -1$. As $\lambda \to \infty$ we have that*

$$\frac{1}{\sqrt{\log \lambda}} \left(M_{a,\beta}(\lambda) - \frac{2\sqrt{\lambda}}{\pi} \right) \Rightarrow \mathcal{N}\Big(0, \frac{1}{\beta \pi^2}\Big).$$

Notice that if one considers the limit on the scale of $(\log \sqrt{\lambda})^{-1/2}$ which in some sense in the natural analogue of the CLT for the Sine_β process, then we actually get convergence to a normal with the same variance as in the Sine_β case. The large deviation result is similarly related to the result in the bulk.

Theorem 2.11. *Fix $\beta > 0$, $a > -1, a \neq -1/2$. The sequence of random variables $\frac{1}{\sqrt{\lambda}} M_{a,\beta}(\lambda)$ satisfies a large deviation principle with scale λ and good rate function $\beta I_{Bess_{a,\beta}}(\rho)$ with*

$$I_{Bess_{a,\beta}}(\rho) = \frac{\nu}{2} + \rho \mathcal{H}(\nu), \qquad \nu = \gamma^{(-1)}(\rho/4), \tag{2.15}$$

where γ is the function given in (2.13).

Remark 2.12. *There is no reason at this point to believe that the restriction that* $a \neq -1/2$ *is necessary. It seems to be a consequence of the proof technique used.*

There is a close relationship between the rate function for the $\mathrm{Bess}_{a,\beta}$ process and the Sine_β process, namely that $I_{\mathrm{Bess}_{a,\beta}}(\rho) = 32 I_{\mathrm{Sine}}(\rho/4)$. As before we can check that the central limit theorem and the LDP are at least formally consistent. Using the relationship between the rate functions we get that $I_{\mathrm{Bess}_{a,\beta}}(\frac{2}{\pi}+x)$ behaves like $-\frac{\pi^2 x^2}{2\log(1/|x|)}$. The substitution $\sqrt{\lambda}x = \sqrt{\frac{1}{\beta\pi^2}\log\lambda} \cdot y$ with a fixed, then the probability would be asymptotically equal to $e^{-y^2/2}$. Lastly, observe that the large deviation result is also consistent with the large gaps result $P(M_{a,\beta}(\lambda) = 0) \sim e^{-\frac{\beta}{2}\lambda}$ [47].

Chapter 3

Point Process Limits of the β-Jacobi Ensemble

In the β- Hermite and Laguerre ensembles discussed above the cases $\beta = 1$, 2 and 4 have the particularity to be solvable and can be studied by means of asymptotics of orthogonal polynomials [44]. In the general β case, the solvability is lost, together with a natural interpretation in terms of classical random matrix ensembles. The soft and hard edge limits for these general ensembles were considered in [49] and [48] respectively, leading to a generalized Tracy-Widom(β) distribution and the stochastic Bessel process, respectively, in the soft edge and the hard edge case. This approach was anticipated in [21], where tridiagonal matrix models associated to the β-Laguerre and Hermite ensemble [17] were conjecture to converge to continuum random operators. It is worth noting that the link between these random operators and the classical results involving the Tracy-Widom distribution remains obscure. To the best of our knowledge, the only result in this direction is [6] where a spiked-random matrix model is investigated.

The object of study in this chapter is the β-Jacobi ensemble (see (1.4)). Note that written in terms of n, n_1, n_2 for large n we will approximate the edges of the $J(n, n_1, n_2, \beta)$ spectrum by

$$\Lambda_{\pm} = \left(\frac{\sqrt{n_1(n_1 + n_2 - n)}}{n_1 + n_2} \pm \frac{\sqrt{nn_2}}{n_1 + n_2} \right)^2. \tag{3.1}$$

This chapter will focus on the behavior at the edge of the spectrum as n, n_1 and n_2 grow to infinity. This approach relies on a tridiagonal representation of the β-Jacobi ensembles [40] and the techniques developed in [49] and [48]. The higher number of parameters makes the phase diagram richer. As there are both a lower and an upper bound on the spectrum, appropriate tuning of the parameters can lead to any combination of soft/hard upper/lower edges. With our notations, the asymptotics of n_1 with respect to n will determine the nature of the lower edge, while respective asymptotics for n_2 will determine the upper edge.

Several works have been devoted to the edge behavior of the β-Jacobi ensembles with varying degree of generality (see for example [16, 19, 15, 45, 46]). The works [11] and [37] are restricted to the cases $\beta = 1$ and 2. Johnstone [37] also provides fluctuation results for the top eigenvalue in the cases $\beta = 1$ and 2. All the aforementioned works treat the case of a single extreme eigenvalue. We note that the work [37] uses the results from [49] and [48] to study very degenerate asymptotics for n_1 and n_2 for which the Jacobi ensemble approximates the Laguerre ensemble.

This work is organized as follows: in Section 3.1.1 we recall the tridiagonal representation for the β-Jacobi ensemble proved in [40] (an alternative approach will be described in the appendix). Proofs in the soft edge cases are presented in Section 3.2 while the hard edge case is treated in Section 3.3. Many details are similar to the corresponding proofs in [49] and [48] and will be omitted.

3.1 Results

3.1.1 Tridiagonal representation

The core of our approach, following [49] and [48], rests heavily on a tridiagonal representation of $J_{n,n_1,n_2,\beta}$. The first tridiagonal representation for the β-Jacobi ensemble was given by Killip and Nenciu [40]. We will work with a somewhat simplified model which was introduced by Sutton ([51], Chapter 5).

Theorem 3.1. *[40] The joint density of the eigenvalues of MM^T is given by $f_{\beta,n,n_1,n_2}(\underline{\lambda})$, where*

$$M = M_{n,n_1,n_2,\beta} = \begin{bmatrix} C_1\tilde{S}_1 & & & & \\ S_2\tilde{C}_1 & C_2\tilde{S}_2 & & & \\ & S_3\tilde{C}_2 & \ddots & & \\ & & \ddots & & \\ & & & S_n\tilde{C}_{n-1} & C_n \end{bmatrix}$$

with $C_i^2 + S_i^2 = 1, \tilde{C}_i^2 + \tilde{S}_i^2 = 1$ and

$$\begin{aligned} C_k &\sim \sqrt{\mathrm{Beta}\left(\frac{\beta}{2}(n_1 - n + k), \frac{\beta}{2}(n_2 - n + k)\right)} \\ \tilde{C}_k &\sim \sqrt{\mathrm{Beta}\left(\frac{\beta}{2}k, \frac{\beta}{2}(n_1 + n_2 - 2n + k + 1)\right)}. \end{aligned}$$

Note that MM^T is indeed tridiagonal. A proof of this theorem is included in the appendix for the sake of completeness.

3.2 Convergence of the spectrum at the soft edge

The proof of the soft edge limit is derived using a more general limiting result by Ramírez, Rider and Virág [49]. This result embeds a sequence of tridiagonal matrices as operators on $L^2[0,\infty)$ and gives conditions for a weak limit under which convergence of the eigenvalues also holds. We use this result to show that the point process $J(n, n_1, n_2, \beta)$ converges to eigenvalues of a random operator.

We begin by stating a general result of Ramírez, Rider, and Virag [49] giving conditions for a sequence of operators to converge to the stochastic Airy operator in an appropriate sense. Our work will then consist of verifying the hypothesis of this theorem in the β-Jacobi case.

Let $H_n : \mathbb{R}^n \to \mathbb{R}^n$ be the linear operator whose associated matrix with respect to the standard bases is symmetric, tridiagonal with diagonal entries $(2m_n^2 + m_n(y_{n,1,k} - y_{n,1,k-1}), k \geq 1)$ and off-diagonal entries $(-m_n^2 + m_n(y_{n,2,k} - y_{n,2,k-1})/2, k \geq 1)$.

We define define the step functions $y_{n,i}(x) = y_{n,i,\lfloor xm_n \rfloor}\mathbf{1}_{xm_n\in[0,n]}$ and make the following assumptions:

Assumption 1 (Tightness/Convergence) There exists a continuous process $x \mapsto y(x)$ such that

$$\begin{aligned} \big(y_{n,i}(x); x \geq 0\big) &\qquad i = 1, 2 \quad \text{are tight in law} \\ \big(y_{n,1}(x) + y_{n,2}(x); x \geq 0\big) &\;\;\Rightarrow\;\; \big(y(x); x \geq 0\big) \quad \text{in law,} \end{aligned}$$

with respect to the Skorokhod topology.

Assumption 2 (Growth/Oscillation bound) There is a decomposition

$$y_{n,i,k} = m_n^{-1} \sum_{\ell=1}^{k} \eta_{n,i,\ell} + \omega_{n,i,k} \tag{3.2}$$

with $\eta_{n,i,k} \geq 0$, such that there are deterministic unbounded nondecreasing continuous functions $\overline{\eta}(x) > 0, \zeta(x) \geq 1$, and random constants $\mu_n(\omega) \geq 1$ defined on the same probability space which satisfy the following: The μ_n are tight in distribution, and, almost surely

$$\overline{\eta}(x)/\mu_n - \mu_n \leq \eta_{n,1}(x) + \eta_{n,2}(x) \leq \mu_n(1 + \overline{\eta}(x)) \tag{3.3}$$

$$\eta_{n,2}(x) \leq 2m_n^2 \tag{3.4}$$

$$|\omega_{n,1}(\xi) - \omega_{n,1}(x)|^2 + |\omega_{n,2}(\xi) - \omega_{n,2}(x)|^2 \leq \mu_n(1 + \overline{\eta}(x)/\zeta(x)) \tag{3.5}$$

for all n and $x, \xi \in [0, m_n]$ with $|x - \xi| \leq 1$.

Theorem 3.2. *[49] Given Assumptions 1 and 2 above and any fixed k, the bottom k eigenvalues of the matrix H_n converge in law to the bottom k eigenvalues of the operator H, where*

$$H = -\frac{d^2}{dx^2} + y'(x).$$

Here the eigenfunction/eigenvalue pairs of H should be understood in the same way as those of $\mathcal{H}_\beta$ as shown in Section 2.1.1. We will show that a tridiagonal matrix with eigenvalues corresponding to the β-Jacobi ensemble satisfies the requirements of the theorem with limiting operator $\mathcal{H}_\beta$.

3.2.1 The β-Jacobi model

We use the following tridiagonal model for the β-Jacobi ensemble. Let

$$Z_{n,\beta} = \begin{bmatrix} C_n & S_n\tilde{C}_{n-1} & & & & \\ & C_{n-1}\tilde{S}_{n-1} & S_{n-1}\tilde{C}_{n-2} & & & \\ & & \ddots & \ddots & & \\ & & & & C_2\tilde{S}_2 & S_1\tilde{C}_2 \\ & & & & & C_1\tilde{S}_1 \end{bmatrix}$$

with $C_k, S_k, \tilde{C}_k$, and $\tilde{S}_k$ defined as in Theorem 3.1. Clearly $\sigma(Z_{n,\beta}Z_{n,\beta}^T) = \sigma(M_{n,\beta}M_{n,\beta}^T)$ in distribution.

Recall the definitions of m_n and α_n from section 2.1.1 given in (2.6). We consider the matrix

$$H_n = \alpha_n\Big((c\tilde{s} + s\tilde{c})^2 I_n - Z_{n,\beta}Z_{n,\beta}^T\Big).$$

This matrix has diagonal and off-diagonal entries given respectively by

$$2m_n^2 + \frac{m_n^2}{cs\tilde{c}\tilde{s}}(c^2\tilde{s}^2 + s^2\tilde{c}^2 - S_{n-k+1}^2\tilde{C}_{n-k}^2 - C_{n-k}^2\tilde{S}_{n-k}^2)$$

and

$$-m_n^2 + \frac{m_n^2}{cs\tilde{c}\tilde{s}}(cs\tilde{c}\tilde{s} - C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1}).$$

Note that the term $(c\tilde{s}+s\tilde{c})^2$ is the Λ_+ defined in (3.1). We define $\Delta y_{n,i,k} = y_{n,i,k} - y_{n,i,k-1}$ with

$$\begin{aligned}
\Delta y_{n,1,k} &= \frac{m_n}{cs\tilde{c}\tilde{s}}(c^2\tilde{s}^2 + s^2\tilde{c}^2 - S^2_{n-k+1}\tilde{C}^2_{n-k} - C^2_{n-k}\tilde{S}^2_{n-k}) \\
\Delta y_{n,2,k} &= \frac{2m_n}{cs\tilde{c}\tilde{s}}(cs\tilde{c}\tilde{s} - C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1}).
\end{aligned}$$

The proof now consists of verifying the hypotheses of Theorem 3.2.

3.2.2 Checking Assumption 1

To show that H_n satisfies Assumption 1 of Theorem 3.2 we use the following proposition which is a simple modification of Theorem 7.4.1 and Corollary 7.4.2 in [26].

Proposition 3.3 ([26])**.** *Let $f \in C^1(\mathbb{R}^+)$ and $g \in C^1(\mathbb{R}^+)$, and let y_n be a sequence of processes with $y_{n,0} = 0$ and independent increments. Assume that*

$$\frac{1}{\epsilon_n}E(\Delta y_{n,k}) = f'(k\epsilon_n) + o(1), \qquad \frac{1}{\epsilon_n}\mathit{Var}\,(\Delta y_{n,k}) = g^2(k\epsilon_n) + o(1), \qquad \frac{1}{\epsilon_n}E(\Delta y_{n,k})^4 = o(1)$$

uniformly for $k\epsilon_n$ on compact sets as $n \to \infty$. Then $y_n(t) = y_{n,\lfloor t/\epsilon_n \rfloor}$ converges in law, with respect to the Skorokhod topology, to the process $f(t) + \int_0^t g(s)db_s$, where b is a standard Brownian motion.

We take $\epsilon_n = 1/m_n$ and apply this to a slightly altered version of $y_{n,1}$ and $y_{n,2}$. Take

$$\begin{aligned}
\Delta \tilde{y}_{n,1,k} &= \frac{m_n}{cs\tilde{c}\tilde{s}}(c^2\tilde{s}^2 + s^2\tilde{c}^2 - S^2_{n-k}\tilde{C}^2_{n-k} - C^2_{n-k}\tilde{S}^2_{n-k}) \\
\Delta \tilde{y}_{n,2,k} &= \frac{2m_n}{cs\tilde{c}\tilde{s}}(cs\tilde{c}\tilde{s} - C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k}).
\end{aligned}$$

Computation gives

$$ES^2_{n-k} = \frac{n_2 - k}{n_1 + n_2 - 2k} = s^2 + \frac{s^2 - c^2}{n_1 + n_2} + \frac{2(n_2 - n_1)}{(n_1 + n_2 - 2\ell)^3}k^2$$

for some $\ell \in [0, k]$. Similar expansions can be written for the other terms involved. We now note that for convergence on compact subsets it is sufficient to consider $k \le cm_n \le c_2 n^{1/3}$ by remark 2.4. Therefore, collecting terms we find that

$$m_n E\Delta\tilde{y}_{n,1,k} = \frac{m_n^2}{cs\tilde{c}\tilde{s}} \cdot \frac{2(c^2 - s^2)(\tilde{c}^2 - \tilde{s}^2)}{n_1 + n_2}\, k + o(1). \tag{3.6}$$

Lemma 3.4. *Let $X = C_{n-k}^i S_{n-k}^j$ with i and j positive integers, then*

$$EX = \sqrt{EX^2} - \frac{1}{8(EX^2)^{3/2}} Var\,(X^2) + O\left(\frac{1}{(n_1+n_2)^2}\right).$$

Similar statements hold for $\tilde{C}_{n-k}$ and $\tilde{S}_{n-k}$.

Here we take $O(1/(n_1 + n_2)^2)$ to mean that there exists a constant C depending only on i and j so that the magnitude of the error is bounded about by $C/(n_1 + n_2)^2$. This lemma will always be used in the event that C_{n-k} or S_{n-k} is raised to an odd power.

Proof. For $a < x \in [0, 1]$

$$\sqrt{a} + \frac{1}{2\sqrt{a}}(x-a) - \frac{1}{8a^{3/2}}(x-a)^2 \le \sqrt{x} \le \sqrt{a} + \frac{1}{2\sqrt{a}}(x-a) - \frac{1}{8a^{3/2}}(x-a)^2 + \frac{(x-a)^3}{16a^{5/2}},$$

and for $x \le a \in [0, 1]$

$$\sqrt{a} + \frac{1}{2\sqrt{a}}(x-a) - \frac{1}{8a^{3/2}}(x-a)^2 + \frac{(x-a)^3}{16a^{5/2}} \le \sqrt{x} \le \sqrt{a} + \frac{1}{2\sqrt{a}}(x-a) - \frac{1}{8a^{3/2}}(x-a)^2.$$

To complete the proof what remains to be shown is that $E(X^2 - EX^2)^3 = O(1/(n_1 + n_2)^2)$. To accomplish this we make the following observation. If $Y \sim \mathrm{Beta}(a, b)$, then for i, j positive integers we have that

$$E(Y^i(1-Y)^j) = \frac{(a+i-1)\cdots(a+1)a\,(b+j-1)\cdots(b+1)b}{(a+b+i+j-1)\cdots(a+b+1)(a+b)}. \tag{3.7}$$

Direct computation then finishes the proof. $\square$

An application of this lemma shows that

$$m_n E\Delta\tilde{y}_{n,2,k} = \frac{m_n^2}{cs\tilde{c}\tilde{s}} \cdot \frac{\tilde{c}^2\tilde{s}^2(c^2-s^2)^2 + c^2s^2(\tilde{c}^2-\tilde{s}^2)^2}{n_1+n_2}k + O\left(\frac{m_n^2}{n_1+n_2}\right) \tag{3.8}$$

Here we mean that for $1 \leq k \leq am_n$ there exists some constant C independent of k, n, n_1, n_2 so that the error is bounded by $Cm_n^2/(n_1+n_2)$. Together with (3.6) this gives us that

$$m_n E(\Delta\tilde{y}_{n,1,k} + \Delta\tilde{y}_{n,2,k}) = \frac{k}{m_n} + o(1) \tag{3.9}$$

which verifies the first condition in Proposition 3.3 with $f(x) = x^2/2$.

The second and fourth moment computations can be done similarly. They give us that

$$m_n E(\Delta\tilde{y}_{n,1,k} + \Delta\tilde{y}_{n,2,k})^2 = \frac{4}{\beta} + o(1), \quad \text{and} \quad m_n E(\Delta\tilde{y}_{n,1,k} + \Delta\tilde{y}_{n,2,k})^4 = o(1). \tag{3.10}$$

Note that, for the fourth moment, it is sufficient to bound $m_n E(\Delta\tilde{y}_{n,2,k})^4$ and (breaking $\tilde{y}_{n,1,k}$ into pieces) to bound the corresponding moments of $s^2\tilde{c}^2 - S_{n-k}^2\tilde{C}_{n-k}^2$ and $c^2\tilde{s}^2 - C_{n-k}^2\tilde{S}_{n-k}^2$. Therefore, Proposition 3.3 can be applied and yields

$$\tilde{y}_{n,1,k} + \tilde{y}_{n,2,k} \Rightarrow \frac{x^2}{2} + \frac{2}{\sqrt{\beta}}b(x),$$

in law in the Skorohod topology.

Recall that in all the above computations, we considered $\tilde{y}$ instead of y. Consequently, all that remains to check the second part of Assumption 1 is to show that $y_{n,1,k} + y_{n,2,k} - \tilde{y}_{n,1,k} - \tilde{y}_{n,2,k}$ converges to the 0 process in law in the Skorohod topology. Direct computation shows that the expectation and variance of the increments are of order $1/n$

and so the expectation and variance of the process go to 0 on compact subsets. This together with a fourth moment bound gives us convergence to the 0 process in law.

Finally, the individual tightness of each $(y_{n,i}(x); x \geq 0)$, $i = 1, 2$ can be obtained along the same lines.

3.2.3 Checking Assumption 2

To check this assumption we again work with the shifted processes $\tilde{y}_{n,i,k}$ and compare this with the original process. To this end we take $\eta_{n,i,k} = m_n E \Delta y_{n,i,k}$, and similarly $\tilde{\eta}_{n,i,k} = m_n E \Delta \tilde{y}_{n,i,k}$. Further, we will neglect the extra -1 found in the $\tilde{C}_{n-k}$ and $\tilde{S}_{n-k}$ terms and show the irrelevance later. We will show the inequality in (3.3) for $\bar{\eta}(x) = x$. Under these definitions we have

$$\tilde{\eta}_{n,1,k} = \frac{m_n^2}{cs\tilde{c}\tilde{s}} \left(c^2\tilde{s}^2 + s^2\tilde{c}^2 - \frac{(n_1-k)(n_1+n_2-n-k)}{(n_1+n_2-2k)^2} - \frac{(n_2-k)(n-k)}{(n_1+n_2-2k)^2} \right)$$
$$\tilde{\eta}_{n,2,k} = \frac{2m_n^2}{cs\tilde{c}\tilde{s}} \left(cs\tilde{c}\tilde{s} - \frac{f(\frac{\beta}{2}(n_1-k))f(\frac{\beta}{2}(n_2-k))f(\frac{\beta}{2}(n-k))f(\frac{\beta}{2}(n_1+n_2-n-k))}{\frac{\beta^2}{4}(n_1+n_2-2k)^2} \right)$$

where $f(x) = \Gamma(x+1/2)/\Gamma(x)$.

We begin with the upper bound on $\tilde{\eta}_{n,1,k} + \tilde{\eta}_{n,2,k}$. The general idea will be to treat $\tilde{\eta}_{n,1,k}$ via a Taylor expansion with a uniform bound on the error term, and work with $\tilde{\eta}_{n,2,k}$ in two regions. The first region will be $1 \leq k \leq \alpha n$ for some $\alpha < 1$ and the second will be $\alpha n \leq k \leq n-1$. On the first region we will again work with primarily with Taylor expansion and find a uniform bound on the error in terms of α, for the second section we return to working with the original $\tilde{\eta}_{n,2,k}$ and show that for $\alpha n \leq k \leq n-1$ this can be bounded by $C_\alpha k/m_n$.

Using the inequality

$$\sqrt{x}\left(1-\frac{2}{x}\right) \le \frac{\Gamma(x+1/2)}{\Gamma(x)} \le \sqrt{x}$$

we get that for $n_1, n_2 > n > k$

$$\begin{aligned}\tilde{\eta}_{n,2,k} \le \frac{2m_n^2}{cs\tilde{c}\tilde{s}}&\left(cs\tilde{c}\tilde{s} - \frac{\sqrt{(n_1-k)(n_2-k)(n-k)(n_1+n_2-n-k)}}{(n_1+n_2-2k)^2}\right)\\ &+15\frac{4m_n^2}{cs\tilde{c}\tilde{s}}\left(\frac{\sqrt{(n_1-k)(n_2-k)(n-k)(n_1+n_2-n-k)}}{\frac{\beta}{2}(n-k)(n_1+n_2-2k)^2}\right).\end{aligned}$$

For convenience we will label the first line of the right hand side by A_k and the second line by B_k. We will treat the second term B_k first.

$$\frac{m_n B_k}{k} \le \frac{120}{\beta}\frac{n_1n_2n(n_1+n_2-n)(n_1+n_2)^3}{n(n_1+n_2-2n)^2(\sqrt{n(n_1+n_2-n)}(n_1-n_2)+\sqrt{n_1n_2}(2n-n_1-n_2))^2}$$

This upper bound has a finite limsup. We now turn to $\tilde{\eta}_{n,1,k}+A_k$ for $1 \le k \le \alpha n$ for some $0<\alpha<1$.

$$\tilde{\eta}_{n,1,k}+A_k = \frac{k}{m_n}+\frac{f(\ell)}{2}k^2$$

for some $\ell \in [1, \alpha n]$. Here $f(\ell)$ is the second derivative of $\tilde{\eta}_{n,1,k}+A_k$ with respect to k. On this range of k we can find explicit upper bounds for

$$m_n k\frac{f(\ell)}{2}$$

in terms of n, n_1, n_2, and α which have finite limsup as n goes to ∞. This together with our bound on $B_k m_n/k$ is enough to give us a sequence of constants c_n so that

$$\tilde{\eta}_{n,1,k}+\tilde{\eta}_{n,2,k} \le c_n\frac{k}{m_n}$$

for $1 \le k \le \alpha n$. For the remaining piece we work with $\tilde{\eta}_{n,1,k}$ and $\tilde{\eta}_{n,2,k}$ separately.

$$\frac{m_n\tilde{\eta}_{n,1,k}}{k} = \frac{m_n^3}{cs\tilde{c}\tilde{s}}\left(\frac{2(n_1-n_2)(2n-n_1-n_2)}{(n_1+n_2)^3}+\frac{6(n_1-n_2)(2n-n_1-n_2)k}{(n_1+n_2-2\ell)^4}\right)$$

for some $1 \leq \ell \leq k \leq n$. Taking the obvious upper bound with $\ell = k = n$ we have an upper bound on $\tilde{\eta}_{n,1,k}$ for $1 \leq k \leq n$ which has a finite limsup. Returning to A_k we note that for $\alpha n \leq k \leq n$ we have that

$$\frac{m_n \tilde{\eta}_{n,2,k}}{k} \leq \frac{m_n^3}{\alpha n}$$

From Remark 2.4 we have that this is an appropriate upper bound. Choosing the larger upper bound on the two sections gives us a sequence μ_n such that

$$\tilde{\eta}_{n,1,k} + \tilde{\eta}_{n,2,k} \leq \mu_n \frac{k}{m_n}$$

for $1 \leq k \leq n-1$.

Turning to the lower bound we have the inequality

$$\begin{aligned}\tilde{\eta}_{n,1,k} + \tilde{\eta}_{n,2,k} \geq & \frac{m_n^2}{cs\tilde{c}\tilde{s}} \left(c^2\tilde{s}^2 + s^2\tilde{c}^2 - \frac{(n_1-k)(n_1+n_2-n-k)}{(n_1+n_2-2k)^2} - \frac{(n_2-k)(n-k)}{(n_1+n_2-2k)^2} \right) \\ & + \frac{2m_n^2}{cs\tilde{c}\tilde{s}} \left(cs\tilde{c}\tilde{s} - \frac{\sqrt{(n_1-k)(n_2-k)(n-k)(n_1+n_2-n-k)}}{(n_1+n_2-2k)^2} \right).\end{aligned} \tag{3.11}$$

One can then make arguments for $1 \leq k \leq \alpha n$ similar to those employed in the upper bound simply by choosing α small enough so that $m_n k f(\ell)/2$ is bounded below by -1. For the remaining k we note that the derivative of the right hand side with respect to k is

$$\frac{m_n^2}{cs\tilde{c}\tilde{s}} \left[\frac{\sqrt{(n-k)(n_1+n_2-n-k)}(n_1-n_2) + \sqrt{(n_1-k)(n_2-k)}(2n-n_1-n_2)}{\sqrt{(n_1-k)(n_2-k)(n-k)(n_1+n_2-n-k)(n_1+n_2)}} \right]^2$$

which is strictly greater than 0 for $1 \leq k \leq n-1$. Therefore

$$\begin{aligned}\tilde{\eta}_{n,1,k}+&\tilde{\eta}_{n,2,k} \geq \\ &\frac{m_n^2}{cs\tilde{c}\tilde{s}}\left(c^2\tilde{s}^2 + s^2\tilde{c}^2 - \frac{(n_1-\alpha n)(n_1+n_2-n-\alpha n)}{(n_1+n_2-2\alpha n)^2} - \frac{(n_2-\alpha n)(n-\alpha n)}{(n_1+n_2-2\alpha n)^2}\right) \\ &\qquad + \frac{2m_n^2}{cs\tilde{c}\tilde{s}}\left(cs\tilde{c}\tilde{s} - \frac{\sqrt{(n_1-\alpha n)(n_2-\alpha n)(n-\alpha n)(n_1+n_2-n-\alpha n)}}{(n_1+n_2-2\alpha n)^2}\right).\end{aligned}$$

This lower bound can be used to get the desired constants for $\alpha n \leq k \leq n-1$ finishing the lower bound. To finish we make the following observation: By direct computation we can find an upper bound δ_n of order $(n_1+n_2)^{-1/3}$ such that

$$|\tilde{\eta}_{n,1,k} + \tilde{\eta}_{n,2,k} - \eta_{n,1,k} - \eta_{n,2,k}| \leq \delta_n.$$

Therefore any error that comes from neglecting the -1 in the $\tilde{C}_{n-k}$ and $\tilde{S}_{n-k}$ terms, or working with the shifted process, may be absorbed into the constant terms. This finishes the proof of (3.3).

We now verify (3.5):

$$\Delta y_{n,2,k} = \frac{2m_n}{cs\tilde{c}\tilde{s}}(cs\tilde{c}\tilde{s} - C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1}) = 2m_n - \frac{2m_n C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1}}{cs\tilde{c}\tilde{s}}.$$

Therefore since all of the relevant pieces are positive we have that

$$m_n E\Delta y_{n,2,k} = 2m_n^2 - E\frac{2m_n^2 C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1}}{cs\tilde{c}\tilde{s}} \leq 2m_n^2.$$

For the proof of the oscillation bound (3.5) we observe that $\sqrt{m_n}\omega_{n,1,k}$ is the sum of 2 martingales. For $\omega_{n,2,k}$ we note that the increments are

$$\begin{aligned}&\frac{2m_n^{3/2}}{cs\tilde{c}\tilde{s}}\left(EC_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1} - C_{n-k}S_{n-k}\tilde{C}_{n-k}\tilde{S}_{n-k-1}\right) \\ &\quad = \frac{2m_n^{3/2}}{cs\tilde{c}\tilde{s}}\Big((E\tilde{S}_{n-k-1} - \tilde{S}_{n-k-1})EC_{n-k}S_{n-k}\tilde{C}_{n-k} \\ &\qquad\qquad + (EC_{n-k}S_{n-k}\tilde{C}_{n-k} - C_{n-k}S_{n-k}\tilde{C}_{n-k})\tilde{S}_{n-k-1}\Big)\end{aligned}$$

and so $\omega_{n,2,k}$ may also be written as the sum of two martingales. Following the corresponding proof in [49] with these martingales gives us the necessary bound. This ends the proof of the convergence at the soft edge.

3.3 Convergence of the spectrum at the hard edge

The proof of the hard edge limit will follow largely from the proof of Ramírez and Rider in [48]. We will formulate a more general theorem which gives conditions for convergence of the eigenvalues, and apply this in the specific setting of the β-Jacobi ensemble. The main idea is again to embed the tridiagonal matrices as operators and show that operator convergence implies convergence of the eigenvalues. However the operator convergence is not shown directly, instead we work with inverse operators to draw our conclusions.

To show the convergence of the spectrum at the hard edge we will first state a more general theorem on convergence of eigenvalues. We will then show that this gives us Theorem 2.7. Though this is stated more generally than the result given by Ramírez and Rider in [48] the majority of the proof follows directly from their work.

3.3.1 A more general setting

Let X_n be a lower bidiagonal matrix

$$X_n = \begin{bmatrix} a_1 & & & & \\ -b_1 & a_2 & & & \\ & -b_2 & a_3 & & \\ & & \ddots & \ddots & \\ & & & -b_{n-1} & a_n \end{bmatrix}$$

We make the following assumptions on the entries:

Assumption 1 There is a Brownian motion $B(\cdot)$, and deterministic functions $r(x)$, $s(x)$ and $\varphi(x)$ continuous on $(0,1)$ such that for $y < x$ in $(0,1]$

$$\frac{n}{a_{\lfloor nx \rfloor}} \Rightarrow r(x) \tag{3.12}$$

$$\sum_{k=\lfloor ny \rfloor}^{\lfloor nx \rfloor} \log\left(\frac{b_k}{a_k}\right) \Rightarrow s(x) - s(y) + \int_y^x \varphi(t) dB_t. \tag{3.13}$$

.

Assumption 2 There exist tight random constants κ_n and κ_n' and $\alpha < 1$ such that

$$\sup_{1 \le k \le n} \frac{E a_k}{a_k} \le \kappa_n \tag{3.14}$$

$$\sum_{k=j}^{i-1} \log\left(\frac{b_k}{a_k}\right) - s(i/n) + s(j/n) \le \kappa_n'(1 + T^\alpha(i/n) + T^\alpha(j/n)) \tag{3.15}$$

Where $T(x) = \int_{1/2}^x \varphi^2(t) dt$.

Assumption 3 Let K_C to be the integral operator with kernel

$$k_C(x,y) = Cr(x) e^{s(x)-s(y)} \exp\left[CT^\alpha(x) + CT^\alpha(y)\right] \mathbf{1}(y < x),$$

then K_C is Hilbert-Schmidt.

Assumption 4 Let K be the integral operator with kernel given by

$$k(x,y) = r(x) e^{s(x)-s(y)} \exp\left[\int_y^x \varphi(t) dB_t\right] \mathbf{1}(y < x).$$

The operator $(KK^T)^{-1}$ has discrete spectrum with simple eigenvalues $0 < \Lambda_0 < \Lambda_1 < \ldots \uparrow \infty$.

Theorem 3.5. *Let X_n be a lower bidiagonal matrix with entries that satisfy Assumptions 1, 2, 3 and 4. Denote the ordered eigenvalues of $X_n X_n^T$ by $\lambda_0 < \lambda < 1 < \ldots$, then*

$$(\lambda_0, \lambda_1, \ldots, \lambda_{k-1}) \Rightarrow (\Lambda_0, \Lambda_1, \ldots, \Lambda_{k-1})$$

jointly in law in the Skorokhod topology for any fixed k as $n \to \infty$.

Remark 3.6. *Assumption 4 can be replaced by further conditions on $r(x)$, $s(x)$ and $\varphi(x)$ in Assumption 1. We omit this here because in our case Assumption 4 can be checked directly.*

Proof. We begin by computing X_n^{-1} and embedding the resulting matrix as an operator on $L^2(0,1]$. By Lemma 4 in [48] we can compute

$$[X_n^{-1}]_{i,j} = \frac{1}{a_i} \prod_{k=j}^{i-1} \frac{b_k}{a_k} \qquad \text{for} \quad j \leq i.$$

The action of our operator on L^2 will then read

$$\left(X_n^{-1} f\right)(x) = \frac{n}{a_{\lfloor nx \rfloor}} \sum_{j=1}^{\lfloor nx \rfloor} \prod_{k=j}^{\lfloor nx \rfloor - 1} \frac{b_k}{a_k} \int_{x_{j-1}}^{x_j} f(x) dx$$

where $x_k = k/n$. We denote this operator by K^n, and note that this is a discrete integral operator with kernel

$$k^n(x,y) = \frac{1}{a_i} \exp\left[\sum_{k=1}^{i-1} \log\left(\frac{b_k}{a_k}\right)\right] \mathbf{1}_L(x,y),$$

where $\mathbf{1}_L(x,y) = \mathbf{1}(x \in [x_{i-1}, x_i))\mathbf{1}(y \in [x_{j-1}, x_j))$ and $i > j$.

To complete the proof we need the following lemmas:

Lemma 3.7. *There exists a probability space on which all K^n and K are defined, and such that any sequence of the operators K^n contains a subsequence which converges to K in Hilbert-Schmidt norm with probability one. In particular for any $n_k \uparrow \infty$ we can find a subsequence $n_{k'} \uparrow \infty$ along which*

$$\lim_{n_{k'} \to \infty} \int_0^1 \int_0^1 |k^{n_{k'}}(x,y)(\omega) - k(x,y)(\omega)|^2 dx dy = 0$$

almost surely.

The remainder of the proof of Theorem 3.5 now follows from the proof of Theorem 1 in [48] with α in place of $3/4$.

$\square$

Proof of Lemma 3.7. We make the following observation: The noise term in the limiting process may be rewritten as

$$\int_x^y \varphi(z) db_z = \tilde{b}\left(T(x)\right) - \tilde{b}\left(T(y)\right).$$

Here the equality is in distribution with a different Brownian motion $\tilde{b}$ living on the same probability space. This can be used to show that limiting operator K is almost surely Hilbert-Schmidt. The remainder of the proof then follows from the proof of Lemma 6 in [48]. The main idea of the proof is to use Skorohod's representation theorem to find a space on which we have almost sure convergence kernels. The tightness conditions then furnish a dominating function which gives L^2 convergence of the kernels as desired.

$\square$

3.3.2 Proof of theorem 2.7

We now apply Theorem 3.5 to our setting. Recall the tridiagonal ensemble of random matrices which spectrum corresponds to the β-Jacobi ensemble. We look at a similar matrix with unchanged spectrum, which can be decomposed into a product of two bidiagonal matrices. Specifically we take

$$W_{n,\beta} = \begin{bmatrix} C_1\tilde{S}_1 & & & & \\ -S_2\tilde{C}_1 & C_2\tilde{S}_2 & & & \\ & -S_3\tilde{C}_2 & \ddots & & \\ & & \ddots & & \\ & & & -S_n\tilde{C}_{n-1} & C_n \end{bmatrix},$$

where the random variables C_k and $\tilde{C}_k$ are again distributed as in Section 3.1.1. Note that $W_{n,\beta}$ is $M_{n,\beta}$ conjugated by $diag(1,-1,1,-1,...)$ and so $\sigma(W_{n,\beta}W_{n,\beta}^T) = \sigma(M_{n,\beta}M_{n,\beta}^T)$. Recall that we are looking for the hard edge limit at the lower edge, therefore we will take $n_1 = n + a_n, a_n \to a \in (-1, \infty)$ with no restriction on n_2 beyond $n_2 \geq n$.

We will apply Theorem 3.5 to the bidiagonal matrix $\sqrt{m_n}W_{n,\beta}$ where $m_n = nn_2$, and then show that the eigenvalue equation $\varphi = \lambda(KK^T)^{-1}\varphi$ is equivalent to $\psi = \lambda\mathfrak{G}_{\beta,a}^{-1}\psi$ completing the proof of Theorem 2.7.

In order to apply Theorem 3.5 we need to show that $\sqrt{m_n}W_{n,\beta}$ satisfies the assumptions. All these assumptions will need to be verified in two cases. The first case will be when $n_2/n \to \gamma \in [1, \infty)$, and the second will be $n_2 \gg n$. Note that these two cases are sufficient because in the general case for any subsequence we can find a further subsequence along which either $n_2/n \to \gamma$ or $n_2 \gg n$ and so convergence of to the eigenvalues of $\mathfrak{G}_{\beta,a}$ in those situations is sufficient. For clarity we note that the discrete kernel of K^n is

$$k_\beta^n(x,y) = \frac{n}{\sqrt{m_n}C_i\tilde{S}_i}\exp\left[\sum_{k=j}^{i-1}\log\left(\frac{S_{k+1}\tilde{C}_k}{C_k\tilde{S}_k}\right)\right]\mathbf{1}_L(x,y),$$

with $\mathbf{1}_L(x,y) = \mathbf{1}_{x\in[x_{i-1},x_i)}\mathbf{1}_{y\in[x_{j-1},x_j)}$.

Verifying Assumption 1

Lemma 3.8. *Assume that $n_2/n \to \gamma \in [1,\infty)$, then for $x \in (0,1]$*

$$\frac{n}{\sqrt{m_n}C_{\lfloor nx \rfloor}\tilde{S}_{\lfloor nx \rfloor}} \Rightarrow \frac{2x+\gamma-1}{\sqrt{\gamma}\sqrt{x(x+\gamma-1)}}.$$

Assume that $n_2 \gg n$ then for $x \in (0,1]$

$$\frac{n}{\sqrt{m_n}C_{\lfloor nx \rfloor}\tilde{S}_{\lfloor nx \rfloor}} \Rightarrow \frac{1}{\sqrt{x}}.$$

Both convergence statements hold in the Skorokhod topology.

Proof. We begin by observing that

$$\text{Var } C_{nx}^2 = O(1/n) \to 0, \quad \text{ as } \quad n \to \infty.$$

And similarly Var $\tilde{S}_{nx}^2 = O(1/n)$, therefore the random variables converge to their expected values in L^2. Now to ensure convergence on the process level fix $\delta > 0$, then for $k > nx$, $x \geq \delta$ we can find a constant c such that

$$E\left(\frac{1}{C_{k+1}} \cdot \frac{\sqrt{a+k+1}}{\sqrt{a+n_2-n+2k+2}} - \frac{1}{C_k} \cdot \frac{\sqrt{a+k}}{\sqrt{a+n_2-n+2k}}\right)^2 \leq \frac{c}{n}$$

and

$$E\left(\frac{1}{\tilde{S}_{k+1}} \cdot \frac{\sqrt{a+n_2-n+k+1}}{\sqrt{a+n_2-n+2k+2}} - \frac{1}{\tilde{S}_k} \cdot \frac{\sqrt{a+n_2-n+k}}{\sqrt{a+n_2-n+2k}}\right)^2 \leq \frac{c}{n}.$$

Kolmogorov's tightness criterion ensures that we have process convergence on the set $[\delta, 1]$. Let $\delta \to 0$ to finish the proof.

□

Lemma 3.9. *Assume that $n_2/n \to \gamma \in [1,\infty)$. There is a Brownian motion $b(\cdot)$ such that for every $y, x \in (0,1]$ with $y < x$ we have*

$$\sum_{k=\lfloor ny \rfloor}^{\lfloor nx \rfloor -1} \log\left(\frac{S_{k+1}\tilde{C}_k}{C_k\tilde{S}_k}\right) \Rightarrow \log\left(\frac{\sqrt{x(\gamma-1+2y)}}{\sqrt{y(\gamma-1+2x)}}\right) + \frac{1+a}{2}\log\left(\frac{y}{x}\right) \tag{3.16}$$
$$+\frac{a}{2}\log\left(\frac{y+\gamma-1}{x+\gamma-1}\right) + \int_y^x \frac{\sqrt{2s+\gamma-1}}{\sqrt{\beta(s^2+\gamma s-s)}}db_s$$

Assume that $n_2 \gg n$. There is a Brownian motion $b(\cdot)$ such that for every $y, x \in (0,1]$ with $y < x$ we have

$$\sum_{k=\lfloor ny \rfloor}^{\lfloor nx \rfloor - 1} \log\left(\frac{S_{k+1}\tilde{C}_k}{C_k \tilde{S}_k}\right) \Rightarrow \frac{a}{2}\log\left(\frac{y}{x}\right) + \frac{1}{\sqrt{\beta}}\int_y^x \frac{db_s}{\sqrt{s}} \tag{3.17}$$

again with both convergence statements holding in the Skorohod topology.

Before beginning the proof we note the following:

Proposition 3.10. *Let $X \sim Beta(p,q)$, then*

$$\mathrm{E}\left(\log\sqrt{\frac{X}{1-X}}\right) = \frac{1}{2}(\Psi_0(p) - \Psi_0(q)), \qquad \mathrm{Var}\left(\log\sqrt{\frac{X}{1-X}}\right) = \frac{1}{4}(\Psi_1(p) + \Psi_1(q)),$$

where Ψ_0 and Ψ_1 are respectively the digamma and trigamma function. Moreover as $x \to \infty$

$$\begin{aligned} \Psi_0(x) &= \frac{\Gamma'(x)}{\Gamma(x)} = \log(x) - \frac{1}{2x} - \frac{1}{12x^2} + O(x^{-4}), \\ \Psi'(x) &= \Psi_0'(x) = \frac{1}{x} + \frac{1}{2x^2} + O(x^{-3}). \end{aligned} \tag{3.18}$$

Proof of Lemma 3.9. We rearrange the process to have independent increments and then apply Proposition 3.3 with $\epsilon_n = 1/n$.

$$\sum_{k=\lfloor ny \rfloor}^{\lfloor nx \rfloor - 1} \log\left(\frac{S_{k+1}\tilde{C}_k}{C_k \tilde{S}_k}\right) = \log C_{\lfloor nx \rfloor} - \log C_{\lfloor ny \rfloor} + \sum_{k=\lfloor ny \rfloor}^{\lfloor nx \rfloor - 1} \log\left(\frac{S_{k+1}\tilde{C}_k}{C_{k+1} \tilde{S}_k}\right).$$

Similar computation to the proof of Lemma 3.8 give us that if $n_2/n \to \gamma$

$$\log\left(\frac{C_{\lfloor nx \rfloor}}{C_{\lfloor ny \rfloor}}\right) \Rightarrow \log\left(\frac{\sqrt{x(\gamma - 1 + 2y)}}{\sqrt{y(\gamma - 1 + 2x)}}\right).$$

And similarly if $n_2 \gg n$

$$\log\left(\frac{C_{\lfloor nx \rfloor}}{C_{\lfloor ny \rfloor}}\right) \Rightarrow \log\left(\frac{\sqrt{x}}{\sqrt{y}}\right).$$

We now consider the process y_n with increments

$$\Delta y_{n,k} = \log(S_{k+1}/C_{k+1}) - \log(\tilde{S}_k/\tilde{C}_k).$$

In the case where $n_2/n \to \gamma$ we find that

$$nE(\Delta y_{n,nx}) = -\frac{1+a}{2x} - \frac{a}{2(x+\gamma-1)} + o(1),$$

and

$$n\mathrm{Var}\ (\Delta y_{n,nx}) = \frac{1}{\beta x} + \frac{1}{\beta(x+\gamma-1)} + o(1).$$

We can also check that $nE(\Delta y_{n,k})^4 = o(1)$. By Proposition 3.3 we get the convergence statement in (3.16).

Similar calculations in the case where $n_2 \gg n$ give you that

$$nE(\Delta y_{n,nx}) = -\frac{1+a}{2x} + o(1), \quad n\mathrm{Var}\ (\Delta y_{n,nx}) = \frac{1}{\beta x} + o(1), \quad nE(\Delta y_{n,k})^4 = o(1).$$

This gives us the convergence in (3.17).

$\square$

Verifying Assumption 2

We now turn to the tightness conditions. For (3.14) we begin by using the sum bound:

$$P\left(\sup_{1\le k\le n} \frac{1}{C_k} \cdot \frac{\sqrt{a_n+k}}{\sqrt{a_n+n_2-n+2k}} > M\right) \le \sum_{k=1}^{n} P\left(C_k^2 < \frac{a_n+k}{M^2(a_n+n_2-n+2k)}\right)$$

Now recall that for $X \sim \Gamma(p,\theta)$, $Y \sim \Gamma(q,\theta)$ we have that $X/(X+Y) \sim \mathrm{Beta}(p,q)$. We then start by finding bound on X and Y before turning to the Beta distribution. In particular we compute $P(X > NE(X))$ and $P(X < E(X)/N)$ for large N: First notice that $E(X) = p\theta$ and so an application of an exponential Chebyshev's inequality with $\theta = 1$ gives us

$$P(X > pN) \leq \left(\frac{2}{e^{N/2}}\right)^p, \qquad \text{and} \qquad P(X < p/N) \leq \left(\frac{e}{1+N}\right)^p$$

Using these bounds we can find an upper bound on

$$P\left(\frac{X}{X+Y} \cdot \frac{p+q}{p} < \frac{1}{N}\right)$$

with exponent $(p \wedge q)$.

When applied to our original setting, this bound is summable in k (as $n \to \infty$), and moreover the resulting sum may be made as small as desired by increasing M. Therefore, for any $\epsilon > 0$ we can choose M such that

$$P\left(\sup_{1\leq k\leq n} \frac{1}{C_k} \cdot \frac{\sqrt{a_n + k}}{\sqrt{a_n + n_2 - n + 2k}} > M\right) < \epsilon,$$

and so the κ_n are tight.

For the tightness of the κ'_n we use the proof of Lemma 5 in [48]. That proof is a reworking of the upper bound in the law of the iterated logarithm and the proof in this situation proceeds with few changes. We note first that this section will again need to be done in two cases. We make the following definitions which are analogous to the A^n_x defined in [49].

$$A^n_x = \sum_{k=j}^{n-1} \left[\log\left(S_{k+1}/C_{k+1}\right) - \log\left(\tilde{S}_k/\tilde{C}_k\right)\right] - \frac{1+a}{2}\log\left(\frac{j}{i}\right) - \frac{a}{2}\log\left(\frac{j+n\gamma-n}{i+n\gamma-n}\right) \tag{3.19}$$

and

$$B^n_x = \sum_{k=j}^{i-1} \left[\log\left(S_{k+1}/C_{k+1}\right) - \log\left(\tilde{S}_k/\tilde{C}_k\right)\right] - \frac{a+1}{2}\log\left(\frac{j}{i}\right). \tag{3.20}$$

for $x \in [x_j, x_{j+1})$. Here A^n_x will be used in the case where $n_2/n \to \gamma < \infty$ and the associated $T(x) = \frac{1}{\beta}\left(\log\frac{1}{x} + \log\frac{1}{\gamma-1+x}\right)$. In the case where $n_2 \gg n$ the object of interest will be B^n_x with $T(x) = \frac{1}{\beta}\log\frac{1}{x}$.

The changes noted above together with the following claim are sufficient to show the κ'_n are tight.

Remark 3.11. *The $T(x)$ as defined in the general theorem differs from the $T(x)$ defined here by a constant. This omission is permissible because we always consider*

$$C(1 + T^{3/4}(x) + T^{3/4}(y))$$

and so the difference may be absorbed by the constant term.

Claim 3.12. *For all $\lambda > 0$ sufficiently small ($\lambda < (\beta/2)[(a+1)\wedge 1]$ will do), if $n_2/n \to \gamma < \infty$ then*

$$E[e^{\lambda A^n_{x_j}}] = \exp\left\{\frac{\lambda^2}{2\beta}\left[\log\left(\frac{1}{x_j}\right) + \log\left(\frac{\gamma}{\gamma - 1 + x_j}\right)\right] + \Theta_n(j)\right\} \tag{3.21}$$

With $|\Theta_n(j)| \le C$ for constant $C = C(a, \beta, \gamma)$. And if $n_2/n \to \infty$

$$E[e^{\lambda B^n_{x_j}}] = \exp\left\{\frac{\lambda^2}{2\beta}\log\left(\frac{1}{x_j}\right) + \Theta_n(j)\right\} \tag{3.22}$$

With $|\Theta_n(j)| \le C$ for constant $C = C(a, \beta)$.

Proof of claim. We have the following products to consider: In the case where $n_2/n \to \gamma$ we have

$$E[e^{\lambda A^n_{x_j}}] = E\prod_{k=j}^{n-1}\left(\frac{S_{k+1}}{C_{k+1}}\right)^{\lambda}\left(\frac{\tilde{C}_k}{\tilde{S}_k}\right)^{\lambda}\left(\frac{k+1}{k}\right)^{\lambda(a+1)/2}\left(\frac{n_2 - n + k + 1}{n_2 - n + k}\right)^{\lambda a/2}, \tag{3.23}$$

and in the case where $n_2 \gg n$

$$E[e^{\lambda B^n_{x_j}}] = E\prod_{k=j}^{n-1}\left(\frac{S_{k+1}}{C_{k+1}}\right)^{\lambda}\left(\frac{\tilde{C}_k}{\tilde{S}_k}\right)^{\lambda}\left(\frac{k+1}{k}\right)^{2(a+1)/2} \tag{3.24}$$

Let's start by considering the portion of (3.23) and (3.24) of the form

$$P = E\prod_{k=j}^{n-1}\left(\frac{S_{k+1}}{C_{k+1}}\right)^{\lambda}\left(\frac{\tilde{C}_k}{\tilde{S}_k}\right)^{\lambda}.$$

We use independence write this as a product of expectations, then by taking logarithms we can consider each term of the resulting sum separately. For the k-th term this gives us $(\log P)_k = I_k + J_k$ where,

$$\begin{aligned} I_k = \; & \log\Gamma\left(\frac{\beta}{2}(n_2 - n + k + 1) + \frac{\lambda}{2}\right) - \log\Gamma\left(\frac{\beta}{2}(n_2 - n + k + 1)\right) \\ & + \log\Gamma\left(\frac{\beta}{2}(a_n + n_2 - n + k + 1) - \frac{\lambda}{2}\right) - \log\Gamma\left(\frac{\beta}{2}(a_n + n_2 - n + k + 1)\right) \end{aligned}$$

and

$$\begin{aligned} J_k = \; & \log\Gamma\left(\frac{\beta}{2}(a_n + k + 1) - \frac{\lambda}{2}\right) - \log\Gamma\left(\frac{\beta}{2}(a_n + k + 1)\right) \\ & + \log\Gamma\left(\frac{\beta}{2}k + \frac{\lambda}{2}\right) - \log\Gamma\left(\frac{\beta}{2}k\right) \end{aligned}$$

From the proof of Claim 10 in [48], we can conclude that

$$I_k = \frac{\lambda^2}{2\beta(n_2 - n + k)} - \frac{\lambda a}{2}\log\left(1 + \frac{1}{n_2 - n + k}\right) + O(1/k^2)$$

and

$$J_k = \frac{\lambda^2}{2\beta k} - \frac{\lambda(a+1)}{2}\log\left(1 + \frac{1}{k}\right) + O(1/k^2).$$

The remaining part of (3.23) gives a contribution of

$$\frac{\lambda a}{2}\log\left(1 + \frac{1}{n_2 - n + k}\right) + \frac{\lambda(a+1)}{2}\log\left(1 + \frac{1}{k}\right).$$

We then use that that

$$\sum_{k=1}^{n} 1/k = \log n + c + O(1/2n)$$

to establish the claim in the case where $n_2/n \to \gamma$. In the case where $n_2 \gg n$, we have that $I_k = O(1/n_2)$ and so may be folded into the constant term. Then the remaining term in (3.24) gives a contribution of

$$\frac{\lambda(a+1)}{2} \log\left(1 + \frac{1}{k}\right)$$

which establishes the claim in this case. □

Verifying Assumption 3

To show that the K_C are Hilbert-Schmidt we have the following proposition:

Proposition 3.13. *For any constant C and $a > -1$, the integral operators on $L^2[0,1]$ with kernels*

$$k_{C,\gamma}(x,y) = C \exp\left[CT^{3/4}(x) + CT^{3/4}(y)\right] \frac{\sqrt{2y+\gamma-1}}{(2x+\gamma-1)^{-1/2}} \frac{(y^2+\gamma y - y)^{a/2}}{(x^2+\gamma x - x)^{(a+1)/2}} \mathbf{1}_{\{y<x\}}$$

where $T(x) = \frac{1}{\beta}\left(\log\frac{1}{x} + \log\frac{1}{\gamma - 1 + x}\right)$ and

$$k_C(x,y) = C \exp\left[C(\log(1/x))^{3/4} + C(\log(1/y))^{3/4}\right] y^{a/2} x^{-(a+1)/2} \mathbf{1}_{\{y<x\}}$$

are Hilbert-Schmidt.

Proof. For the operator with kernel $k_{C,\gamma}$ the change of variable $x^2 + \gamma x - x = e^{-s}$ and $y^2 + \gamma y - y = e^{-t}$ gives the square of the Hilbert-Schmidt norm

$$\int_0^1 \int_0^1 |k_{C,\gamma}(x,y)|^2 dx dy = C^2 \int_0^\infty e^{2Cs^{3/4} + as} \int_s^\infty e^{2Ct^{3/4} - (a+1)t} dt ds.$$

Similarly for the operator with kernel $k_C(x,y)$ we can make the change of variables $x = e^{-s}$ and $y = e^{-t}$ to find the same double integral. This operator is finite if and only if $a > -1$, so both operators are Hilbert-Schmidt. □

Verifying Assumption 4

To show that the eigenvalues of $(KK^T)^{-1}$ are simple with $0 < \Lambda_0 < \Lambda_1 < \cdots \uparrow \infty$ we make the following observations:

Observation 1: In the case where $n_2/n \to \gamma$ the spectral problem reads

$$f(x) = \lambda K^T K f(x) = \lambda \int_0^1 k(y,x) \int_0^1 k(y,z) f(z) dz dy \tag{3.25}$$

$$f(x) = \frac{\lambda}{\gamma} \int_x^1 \frac{\sqrt{2x+\gamma-1}}{(x^2+\gamma x - x)^{-a/2}} \frac{2y+\gamma-1}{(y^2+\gamma y - y)^{a+1}} \exp\left(\int_x^y \sqrt{\frac{2s+\gamma-1}{\beta s(s+\gamma-1)}} db_s \right)$$
$$\times \int_0^y \frac{\sqrt{2z+\gamma-1}}{(z^2+\gamma z-1)^{-a/2}} \exp\left(\int_z^y \sqrt{\frac{2s+\gamma-1}{\beta s(s+\gamma-1)}} db_s \right) f(z) dz dy.$$

With a bit of rearranging we find that for

$$g(x) = \frac{(x^2+\gamma x - x)^{-a/2}}{\sqrt{2x+\gamma-1}} \exp\left(- \int_x^1 \sqrt{\frac{2s+\gamma-1}{\beta s(s+\gamma-1)}} db_s \right) f(x),$$

under the change of variables $(x(x+\gamma-1), y(y+\gamma-1), z(z+\gamma-1)) \mapsto (\gamma p, \gamma q, \gamma r)$ with $h(p) = g(x)$ the previous equation reads

$$h(p) = \lambda \int_0^1 r^a e^{\frac{2}{\sqrt{\beta}} \hat{b}(\log 1/r)} h(r) \int_{p \vee r}^1 q^{-(a+1)} e^{-\frac{2}{\sqrt{\beta}} \hat{b}(\log 1/q)} dq dr. \tag{3.26}$$

Observation 2: In the case where $n_2 \gg n$ the spectral problem instead reads as

$$f(x) = \lambda \int_x^1 x^{a/2} y^{-a-1} \exp\left(\frac{1}{\sqrt{\beta}} \int_x^y \frac{db_s}{\sqrt{s}} \right) \int_0^y z^{a/2} \exp\left(\frac{1}{\sqrt{\beta}} \int_z^y \frac{db_s}{\sqrt{s}} \right) f(z) dz dy.$$

We then take

$$g(x) = x^{-a/2} \exp\left(\frac{-1}{\sqrt{\beta}} \int_x^1 \frac{db_s}{\sqrt{s}} \right) f(x)$$

which again yields the equation

$$g(x) = \lambda \int_0^1 z^a e^{\frac{2}{\sqrt{\beta}} \hat{b}(\log 1/z)} \int_{x \vee z}^1 y^{-(a+1)} e^{-\frac{2}{\sqrt{\beta}} \hat{b}(\log 1/y)} g(z) dy dz. \tag{3.27}$$

Equations (3.26) and (3.27) are equivalent to the eigenvalue equation $\psi = \lambda \mathfrak{G}_{\beta,a}\psi$ ([48], Proof of Theorem1). The simplicity of the eigenvalues of $\mathfrak{G}_{\beta,a}$ and their ordering with $0 < \Lambda_0(\beta, a) < \Lambda_1(\beta, a) < \cdots \uparrow \infty$ are discussed by Ramírez and Rider in [48].

We have showed that Theorem 3.5 applies to $W_{n,\beta}$ and the spectrum of the limiting operator is the same as that of $\mathfrak{G}_{\beta,a}$. This completes the proof of Theorem 2.7.

Chapter 4

Large Deviations for the Sine$_\beta$ Process

Before the proof of the large deviation of the Sine$_\beta$ process we will first prove a similar result for a simpler related process. This process arises as a limit of a related symmetric random matrix ensemble. Let $H_{n,\sigma}$ be a random symmetric tridiagonal matrix with entries equal to 1 above and below the diagonal and i.i.d. normals with mean zero and variance $\frac{\sigma^2}{n}$ on the diagonal.

$$H_{n,\sigma} = \begin{pmatrix} \omega_1 & 1 & & & \\ 1 & \omega_2 & 1 & & \\ & 1 & \ddots & & \\ & & & \ddots & 1 \\ & & & 1 & \omega_n \end{pmatrix}, \qquad \omega_i \sim N(0, \sigma^2 n^{-1}). \tag{4.1}$$

The matrix $H_{n,\sigma}$ can be viewed as a one-dimensional discrete random Schrödinger operator. In [43] it was shown that the bulk scaling limit of the spectrum of $H_{n,\sigma}$ (along appropriate subsequences) is a point process with density $(2\pi)^{-1}$ denoted by Sch_τ. (The parameter $\tau > 0$ depends on σ and the point in the spectrum where we zoom in to take the limit.) The process Sch_τ can be characterized via its counting function in a similar way to the Sine_β process. Consider the following one-parameter family of SDEs:

$$d\phi_\lambda = \lambda dt + dB_0 + \mathrm{Re}\left[e^{-i\phi\lambda}(dB_1 + idB_2)\right], \qquad \phi_\lambda(0) = 0, \quad t \in [0, \infty) \tag{4.2}$$

where B_0, B_1, B_2 are independent standard Brownian motions. Then the random set

$$\Lambda_\tau := \{\lambda : \phi_{\lambda/\tau}(\tau) \in 2\pi\mathbb{Z}\}$$

has the same distribution as Sch_τ. Denote the counting function of the process by $\widetilde{N}_\tau$, i.e. for $\lambda > 0$ let $\widetilde{N}_\tau(\lambda) = \#(\mathrm{Sch}_\tau \cap [0, \lambda])$. In [43] it was shown that $\widetilde{N}_\tau(\lambda)$ is close to a normal with mean $\frac{\lambda}{2\pi}$ and a constant variance $\frac{\tau}{4\pi^2}$. In our next result we derive the large deviation behavior of $\widetilde{N}_\tau(\lambda)$, this is the analogue of Theorem 2.9 for the Sch_τ processes.

Theorem 4.1. *Fix $\tau > 0$. The sequence of random variables $\frac{1}{\lambda}\widetilde{N}_\tau(\lambda)$ satisfies a large deviation principle with scale λ^2 and rate function $\frac{1}{\tau}I_{\mathrm{Sch}}(\cdot)$ where $I_{\mathrm{Sch}}(\rho) = \mathcal{I}(2\pi\rho)$ and for $q > 0$*

$$\mathcal{I}(q) = \frac{2-a}{8} - \frac{E(a)}{4K(a)}, \quad \text{with} \quad a = a(q) = K^{-1}(\pi/(2q)) \tag{4.3}$$

and $\mathcal{I}(0) = 1/8$.

The rate function $I_{\mathrm{Sch}}(\rho)$ is strictly convex and locally quadratic at the absolute minimum point $\rho = \frac{1}{2\pi}$. (See Proposition B.1.) The local behavior of $I_{\mathrm{Sch}}(\rho)$ at $\rho = \frac{1}{2\pi}$ is formally consistent with the fact that $N_\tau(\lambda) - \frac{\lambda}{2\pi}$ is close to a normal random variable with a constant variance $\frac{\tau}{4\pi^2}$.

The proofs of Theorems 2.9 and 4.1 will rely on path level large deviation principles on the corresponding stochastic differential equations. These in turn will follow by analyzing the hitting time of 2π for the diffusion

$$d\widetilde{\alpha}_\lambda = \lambda dt + 2\sin(\widetilde{\alpha}_\lambda/2)dB, \qquad \widetilde{\alpha}_\lambda(0) = 0, \quad t \in [0, \infty). \tag{4.4}$$

Note, that for a fixed λ the process $\widetilde{\alpha}_\lambda(t)$ is equal in distribution to $\phi_\lambda(t) - \phi_0(t)$ from (4.2).

In the next section we summarize some of the important properties of the SDEs we work with, and state the needed path level large deviation results. In Section 4.2 we study diffusion $\widetilde{\alpha}_\lambda$ of (4.4) using the Cameron-Martin-Girsanov change of measure technique. In Sections 4.3 and 4.4 we derive path level large deviations for the diffusions α_λ and $\widetilde{\alpha}_\lambda$ from (2.9) and (4.4). In Section 4.5 we analyze the rate functions for the path level large deviations and in Section 4.6 we complete the proofs of Theorems 2.9 and 4.1. In the Appendix we will discuss various properties and asymptotics of the used special functions.

4.1 Properties of the diffusions corresponding to Sine$_\beta$ and Sch$_\tau$

Our starting point is the observation that if $\lambda > 0$ is fixed, then if the diffusion $\widetilde{\alpha}_\lambda$ (defined in (4.4)) hits $2n\pi$ for $n \in \mathbb{Z}$, it will stay above it. This can be seen from the fact that when $\widetilde{\alpha}_\lambda$ hits $2n\pi$ the noise term vanishes, but the drift term is always positive. Introduce the notations

$$\lfloor y \rfloor_{2\pi} = \max\{2\pi k : 2\pi k \le y\}, \qquad \lceil y \rceil_{2\pi} = \min\{2\pi k : 2\pi k \ge y\}.$$

From the strong Markov property we immediately get the following proposition.

Proposition 4.2. *Fix $\lambda > 0$. Then the process $\lfloor \widetilde{\alpha}_\lambda(t) \rfloor_{2\pi}$ is non-decreasing in t. Moreover, the waiting times between the jump times of this process are i.i.d. with the same* *distribution as the hitting time*

$$\tau_\lambda = \inf\{t : \widetilde{\alpha}_\lambda(t) \geq 2\pi\}. \tag{4.5}$$

Consider the diffusions $\widetilde{\alpha}_\lambda^{(1)}$ and $\widetilde{\alpha}_\lambda^{(2)}$ which are strong solutions of the SDE (4.4), but with initial conditions $\widetilde{\alpha}_\lambda^{(1)}(0) = c_1 \leq \widetilde{\alpha}_\lambda^{(2)} = c_2$. Then a simple coupling argument shows that $\widetilde{\alpha}_\lambda^{(1)}(t) \leq \widetilde{\alpha}_\lambda^{(2)}(t)$ for all $t \geq 0$. Our next proposition will build on this statement using the strong Markov property.

Proposition 4.3. *Let $0 = t_0 < t_1 < t_2 < \cdots < t_n = T$ and fix a $\lambda > 0$. Consider the solution $\widetilde{\alpha}_\lambda(t)$ of (4.4) on $[0, T]$. Then there exists independent random variables $\xi_1, \xi_2, \ldots, \xi_n$ so that*

$$\lfloor \xi_i \rfloor_{2\pi} \leq \lfloor \widetilde{\alpha}_\lambda(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi} \leq \lfloor \xi_i \rfloor_{2\pi} + 2\pi, \quad 1 \leq i \leq n, \tag{4.6}$$

and ξ_i is distributed as $\widetilde{\alpha}_\lambda(t_i - t_{i-1})$.

Proof. Let $\hat{\alpha}_i(s)$ be defined as the strong solution of (4.4) on $[t_{i-1}, t_i]$ with initial condition $\hat{\alpha}_i(t_{i-1}) = 0$ and let $\xi_i = \hat{\alpha}_i(t_i)$. Clearly, $\xi_i, 1 \leq i \leq n$ are independent random variables and $\xi_i \stackrel{d}{=} \widetilde{\alpha}_\lambda(t_i - t_{i-1})$, we just have to show that (4.6) holds. Fix an integer $1 \leq i \leq n$ and define

$$\widetilde{\alpha}_\lambda^{(1)}(s) = \hat{\alpha}_i(s) + \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi}, \qquad \widetilde{\alpha}_\lambda^{(2)}(s) = \hat{\alpha}_i(s) + \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi} + 2\pi, \qquad s \in [t_{i-1}, t_i].$$

Then $\widetilde{\alpha}_\lambda, \widetilde{\alpha}_\lambda^{(1)}, \widetilde{\alpha}_\lambda^{(2)}$ are all strong solutions of (4.4) on $[t_{i-1}, t_i]$ with initial conditions

$$\widetilde{\alpha}_\lambda^{(1)}(t_{i-1}) \leq \widetilde{\alpha}_\lambda(t_{i-1}) \leq \widetilde{\alpha}_\lambda^{(2)}(t_{i-1}) = \widetilde{\alpha}_\lambda^{(1)}(t_{i-1}) + 2\pi.$$

The ordering is preserved by the coupling so we have

$$\widetilde{\alpha}_\lambda^{(1)}(t_i) \leq \widetilde{\alpha}_\lambda(t_i) \leq \widetilde{\alpha}_\lambda^{(2)}(t_i) = \widetilde{\alpha}_\lambda^{(1)}(t_i) + 2\pi.$$

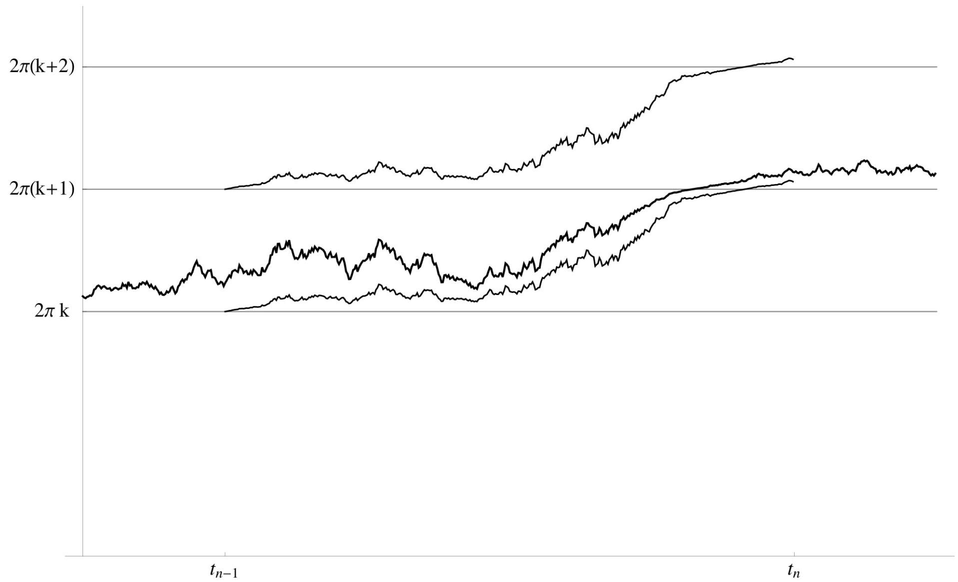

Figure 4.2: The coupling of Proposition 4.3. The process $\widetilde{\alpha}_\lambda$ is the diffusion in the middle, it is sandwiched between $\widetilde{\alpha}_\lambda^{(1)}$ and $\widetilde{\alpha}_\lambda^{(2)} = \widetilde{\alpha}_\lambda^{(1)} + 2\pi$ which start at integer multiples of 2π at the beginning of the coupling interval.

(See Figure 4.2 for an illustration.) From this we get

$$\begin{aligned}\lfloor \widetilde{\alpha}_\lambda(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi} &= \lfloor \widetilde{\alpha}_\lambda(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda^{(1)}(t_{i-1}) \rfloor_{2\pi} \\ &\geq \lfloor \widetilde{\alpha}_\lambda^{(1)}(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda^{(1)}(t_{i-1}) \rfloor_{2\pi} = \lfloor \xi_i \rfloor_{2\pi},\end{aligned}$$

and

$$\begin{aligned}\lfloor \widetilde{\alpha}_\lambda(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi} &= \lfloor \widetilde{\alpha}_\lambda(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda^{(1)}(t_{i-1}) \rfloor_{2\pi} \\ &\leq \lfloor \widetilde{\alpha}_\lambda^{(2)}(t_i) \rfloor_{2\pi} - \lfloor \widetilde{\alpha}_\lambda^{(1)}(t_{i-1}) \rfloor_{2\pi} = \lfloor \xi_i \rfloor_{2\pi} + 2\pi. \qquad \square\end{aligned}$$

We will also need another type of coupling for a slightly more general family of diffusions. Consider the SDE

$$d\xi_{f,c} = f dt + \mathrm{Re}((e^{-i\xi_{f,c}} - 1)(dB_1 + i dB_2)), \quad \xi_{f,c}(0) = c, \quad t \in [0, \infty) \tag{4.7}$$

where f is an integrable non-negative function. Note, that for fixed f, c this process has the same distribution as

$$d\widetilde{\xi}_{f,c} = f dt + 2\sin(\widetilde{\xi}_{f,c}/2)dB, \quad \widetilde{\xi}_{f,c}(0) = c, \quad t \in [0,\infty) \tag{4.8}$$

The following properties of $\xi_{f,c}$ follow from the basic theory of diffusions and standard coupling arguments.

Proposition 4.4. *(i) Let $\tau_{2\pi n}$ be the hitting time of $2\pi n$, where $2\pi n > c$ and n is an integer. Then for any $t > \tau_{2\pi n}$ we have $\xi_{f,c} \geq 2\pi n$. In particular, if $c \geq 0$ then $\xi_{f,c}(t)$ stays non-negative for all $t > 0$.*

(ii) If $f \geq g$ and $\xi_{f,a}$ and $\xi_{g,b}$ are driven by the same Brownian motions then $\xi_{f,a} - \xi_{g,b}$ has the same distribution as $\xi_{f-g,a-b}$. If $a \geq b$ then $\xi_{f,a} - \xi_{g,b}$ stays a.s. non-negative for all t.

(iii) For any finite T we have the following exponential tail bound

$$P(\xi_{f,0}(T) \geq ka) \leq 2\left(\frac{\int_0^T f(t)dt}{2\pi a}\right)^k, \qquad k \in \mathbb{N}. \tag{4.9}$$

If $\int_0^\infty f(t)dt < \infty$ then $\xi_{f,c}(\infty) = \lim_{t\to\infty} \xi_{f,c}(t)$ exists a.s. and the previous bound holds for $T = \infty$ as well.

Sketch of the proof. The first statement follows from the strong Markov property and the fact that in (4.8) the noise term vanishes if $\widetilde{\xi}_{f,c} \in 2\pi\mathbb{Z}$, but the drift is always non-negative. The first part of (ii) follows by considering the difference of the SDEs for $\xi_{f,a}$, $\xi_{g,b}$ and noting that $(e^{-i\xi_{f,a}} - e^{-i\xi_{g,b}})(dB_1 + idB_2)$ has the same distribution as $(e^{-i(\xi_{f,a}-\xi_{g,b})} - 1)(dB_1 + idB_2)$. The second part of (ii) follows from the first statement. Finally, (4.9) follows from the Markov inequality and the strong Markov property. The existence of the limit is proved in Proposition 9 of [54]. (See that proposition for more details on the proof.) □

Our main theorems will follow from the following path level large deviations.

Theorem 4.5. *Fix $\beta > 0$ and let $\alpha_\lambda(t)$ be the process defined in (2.8) or (2.9). Then the sequence of rescaled processes $(\frac{\alpha_\lambda(t)}{\lambda}, t \in [0,\infty))$ satisfies a large deviation principle on $C[0,\infty)$ with the uniform topology with scale λ^2 and good rate function $\mathcal{J}_{\mathrm{Sine}_\beta}$. The rate function $\mathcal{J}_{\mathrm{Sine}_\beta}$ is defined as*

$$\mathcal{J}_{\mathrm{Sine}_\beta}(g) = \int_0^\infty \mathfrak{f}^2(t)\mathcal{I}\left(g'(t)/\mathfrak{f}(t)\right) dt, \quad \text{with} \quad \mathfrak{f}(t) = \mathfrak{f}_\beta(t) = \tfrac{\beta}{4} e^{-\frac{\beta}{4}t}$$

in the case where $g(0) = 0$ and g is absolutely continuous with non-negative derivative g'. In all other cases $\mathcal{J}_{\mathrm{Sine}_\beta}(g)$ is defined as ∞.

Theorem 4.6. *Fix $T > 0$ and let $\widetilde{\alpha}_\lambda(t)$ be the process defined in (4.4). Then the sequence of rescaled processes $(\frac{\widetilde{\alpha}_\lambda(t)}{\lambda}, t \in [0,T])$ satisfies a large deviation principle on $C[0,T]$ with scale λ^2 and good rate function $\mathcal{J}_{\mathrm{Sch},T}$. The rate function is defined as*

$$\mathcal{J}_{\mathrm{Sch},T}(g) = \int_0^T \mathcal{I}\left(g'(t)\right) dt$$

in the case where $g(0) = 0$ and g is absolutely continuous with non-negative derivative g', and $\mathcal{J}_{\mathrm{Sch},T}(g) = \infty$ in all other cases.

In order to prove Theorem 4.6 we observe that $\frac{\widetilde{\alpha}_\lambda(t)}{\lambda}$ is close to $\frac{\lfloor \widetilde{\alpha}_\lambda(t) \rfloor_{2\pi}}{\lambda}$ for large λ and by Proposition 4.2 we only need to analyze the hitting time τ_λ to understand the evolution of $\lfloor \widetilde{\alpha}_\lambda(t) \rfloor_{2\pi}$. The proof of Theorem 4.5 will follow along similar lines after approximating the drift in (2.8) with a piecewise constant function.

4.2 Analysis of the hitting time τ_λ

The following proposition summarizes our bounds on the relevant hitting times.

Proposition 4.7. *Let $\tau_\lambda = \inf\{t : \widetilde{\alpha}_\lambda(t) \geq 2\pi\}$ where $\widetilde{\alpha}_\lambda$ is the solution of (4.4) and fix $a < 1$. Then we have*

$$E e^{\frac{\lambda^2 a}{8}\tau_\lambda - \frac{\lambda(|a| \wedge \sqrt{|a|})}{4}\tau_\lambda} \leq e^{-\lambda \mathcal{H}(a)}. \tag{4.10}$$

Let $t_a = 4K(a)$ and fix $0 < \varepsilon < |t_a - 2\pi|$. Then we have

$$P(\lambda\tau_\lambda \in [t_a - \varepsilon, t_a + \varepsilon]) \geq A(\varepsilon, \lambda, a) e^{-\lambda(\mathcal{H}(a) + \frac{a t_a}{8}) - \lambda\frac{|a|\varepsilon}{8} - \lambda\frac{|a|}{2}(t_a + \varepsilon)} \tag{4.11}$$

where $\lim_{\lambda\to\infty} A(\varepsilon, \lambda, a) = 1$ for fixed a, ε.

Our first step is a change of variables in (4.4). We introduce $X_\lambda(t) = \log(\tan(\widetilde{\alpha}_\lambda(t)/4))$, by Itô's formula this satisfies the SDE

$$dX_\lambda = \frac{\lambda}{2}\cosh X_\lambda\, dt + \frac{1}{2}\tanh X_\lambda\, dt + dB_t, \qquad X_\lambda(0) = -\infty. \tag{4.12}$$

The distribution of the hitting time of 2π for $\widetilde{\alpha}_\lambda(t)$ is the same as that of the hitting time of ∞ for X_λ. With a small abuse of notation from now on we will use the notation τ_λ for the blow-up time of $X_\lambda(t)$, i.e. $\tau_\lambda = \sup\{t : X_\lambda(t) < \infty\}$. In order to study τ_λ we will introduce a similar diffusion with a modified drift. Let $a < 1$ and consider

$$dY_{\lambda,a} = \frac{\lambda}{2}\sqrt{\cosh^2 Y_{\lambda,a} - a}\, dt + \frac{1}{2}\tanh Y_{\lambda,a} dt + dB_t, \qquad Y_{\lambda,a}(0) = -\infty. \tag{4.13}$$

To prove Proposition 4.7 we will choose an appropriate a and compare X_λ with the diffusion $Y_{\lambda,a}$ using the Cameron-Martin-Girsanov formula. Introduce the following notations for the drifts:

$$f_\lambda(x) = \frac{\lambda}{2}\cosh x + \frac{1}{2}\tanh x, \qquad h_{\lambda,a}(y) = \frac{\lambda}{2}\sqrt{\cosh^2 y - a} + \frac{1}{2}\tanh y.$$

Note, that we have the uniform bound

$$|f_\lambda(x) - h_{\lambda,a}(x)| \leq \tfrac{1}{2}\lambda|a|. \tag{4.14}$$

The following proposition will be our main tool for our estimates.

Proposition 4.8. *Fix $a < 1$ and consider $X = X_\lambda$ and $Y = Y_{\lambda,a}$. Denote by τ_λ and $\tau_{Y,\lambda}$ the blowup times of X and Y. Then for any $s > 0$ we have*

$$P(\lambda\tau_\lambda > s) = E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} > s)e^{-G_{s/\lambda}(Y)}\right], \tag{4.15}$$

and where τ denotes the explosion of the appropriate diffusion to infinity

$$1 = Ee^{-G_{\tau\wedge(s/\lambda)}(Y)} = Ee^{G_{\tau\wedge(s/\lambda)}(X)}, \tag{4.16}$$

where

$$G_s(X) = \int_0^s h_{\lambda,a}(X(t)) - f_\lambda(X(t))dX - \frac{1}{2}\int_0^s (h^2_{\lambda,a}(X) - f^2_\lambda(X))dt.$$

Proof. This is just the Cameron-Martin-Girsanov formula for diffusions with explosion. Note, that because of (4.14) the process $e^{G_{\tau\wedge s}(X)}$ satisfies the Novikov criterion and it is a positive martingale. From this the usual steps of the proof can be completed (see e.g. [39], [38]). □

Proof of Proposition 4.7. We first estimate the Girsanov exponent

$$\begin{aligned} G_s(X) = \frac{\lambda}{2}\int_0^s &(\sqrt{\cosh^2 X - a} - \cosh X)dX \\ &- \frac{1}{2}\int_0^s \left(-\frac{\lambda^2}{4}a + \frac{\lambda}{2}(\sqrt{\cosh^2 X - a} - \cosh X)\tanh X\right)\, dt. \end{aligned}$$

Applying Itô's formula for $\theta(X) = h_{\lambda,a}(X) - f_\lambda(X)$ we have that $\int_0^t \theta(X)dX = \int_{X_0}^{X_t} \theta(x)dx - \frac{1}{2}\int_0^s \theta'(X)dt$. This gives us

$$G_s(X) = \frac{\lambda^2 a}{8}s + \frac{\lambda}{2}\int_{-\infty}^{X_s}(\sqrt{\cosh^2 x - a} - \cosh x)dx$$
$$+ \frac{\lambda}{4}\int_0^s \frac{a\tanh}{\sqrt{\cosh^2 X - a}} \cdot \frac{\sqrt{\cosh^2 X - a} - \cosh X}{\sqrt{\cosh^2 X - a} + \cosh X} ds.$$

Note, that

$$\frac{1}{2}\int_{\mathbb{R}}(\sqrt{\cosh^2 x - a} - \cosh x)dx = -\int_0^{\pi/2} \frac{a}{1 + \sqrt{1 - a\sin^2 y}} dy$$
$$= (1-a)K(a) - E(a) = \mathcal{H}(a),$$

where this last equality can be seen by differentiating both sides with respect to a and checking equality at $a = 0$. It is not hard to check that

$$\left|\frac{a\tanh}{\sqrt{\cosh^2 X - a}} \frac{\sqrt{\cosh^2 X - a} - \cosh X}{\sqrt{\cosh^2 X - a} + \cosh X}\right| \leq \left|\frac{a\tanh x}{\sqrt{\cosh^2 x - a}}\right| \leq |a| \wedge \sqrt{|a|}, \qquad \text{for} \quad a < 1$$

uniformly in x. The upper bound $|a|$ follows from $\sqrt{\cosh^2 x - a} \geq |\sinh x|$, while the bound $\sqrt{|a|}$ requires the optimization of the function $\frac{|a|\sqrt{y-1}}{\sqrt{y}\sqrt{y-a}}$ for $y \geq 1$. This gives the bound

$$\left|G_{\tau_\lambda}(X) - \frac{\lambda^2 a\tau}{8} - \lambda\mathcal{H}(a)\right| \leq \frac{\lambda\tau(|a| \wedge \sqrt{|a|})}{4}. \tag{4.17}$$

To get the exponential moment bound (4.10) we use $1 = Ee^{G_{\tau\wedge s/\lambda}(X)}$ from (4.16). We let $s \to \infty$, use Fatou's lemma and (4.17) to get

$$1 \geq Ee^{G_\tau(X)} \geq Ee^{\frac{\lambda^2 a}{8}\tau + \lambda\mathcal{H}(a) - \frac{\lambda(|a| \wedge \sqrt{|a|})\tau}{4}}. \tag{4.18}$$

Rearranging the terms we get (4.10).

To prove the lower bound (4.11) we write

$$
\begin{aligned}
P(\lambda\tau_\lambda \in (t_a-\varepsilon, t_a+\varepsilon)) &= P(\lambda\tau_{Y,\lambda} > t_a-\varepsilon) - P(\lambda\tau_{Y,\lambda} > t_a+\varepsilon)\\
&= E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} > t_a-\varepsilon)e^{-G_{\tau\wedge(t_a-\varepsilon)/\lambda}(Y)}\right] - E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} > t_a+\varepsilon)e^{-G_{\tau\wedge(t_a+\varepsilon)/\lambda}(Y)}\right]\\
&= E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} \in (t_a-\varepsilon, t_a+\varepsilon))e^{-G_{\tau\wedge(t_a+\varepsilon)/\lambda}(Y)}\right]\\
&= E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} \in (t_a-\varepsilon, t_a+\varepsilon))e^{-G_{\tau}(Y)}\right], \qquad (4.19)
\end{aligned}
$$

where we used the fact that $e^{-G_{\tau\wedge t}(Y)}$ is martingale in the third line. Because of (4.17) we have

$$
G_\tau(Y) \le \frac{\lambda^2 a\tau}{8} + \lambda\mathcal{H}(a) + \frac{\lambda|a|\tau}{4}, \qquad (4.20)
$$

and we can bound the last expectation as

$$
\begin{aligned}
E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} \in (t_a-\varepsilon, t_a+\varepsilon))e^{-G_{\tau}(Y)}\right] &\ge E\left[\mathbf{1}(\lambda\tau_{Y,\lambda} \in (t_a-\varepsilon, t_a+\varepsilon))e^{-\frac{\lambda^2 a\tau}{8}-\lambda\mathcal{H}(a)-\frac{\lambda|a|\tau}{4}}\right]\\
&\ge P(\lambda\tau_{Y,\lambda} \in (t_a-\varepsilon, t_a+\varepsilon))e^{-\frac{\lambda a(t_a\pm\varepsilon)}{8}-\lambda\mathcal{H}(a)-\frac{\lambda|a|(t_a+\varepsilon)}{4}},
\end{aligned}
$$

where we choose the sign of ε in $t_a \pm \varepsilon$ the same way as the sign of a.

If we can show that $\lim_{\lambda\to\infty} P(\lambda\tau_{Y,\lambda} \in (t_a-\varepsilon, t_a+\varepsilon)) = 1$ for fixed a and ε then this will complete the proof of (4.11). Note, that $\widetilde{Y}(t) := Y_{\lambda,a}(t/\lambda)$ satisfies the SDE

$$
d\widetilde{Y} = \frac{1}{2}\sqrt{\cosh^2 Y - a}dt + \frac{1}{2\lambda}\tanh\widetilde{Y}dt + \frac{1}{\sqrt{\lambda}}dB_t, \qquad \widetilde{Y}(0) = -\infty.
$$

As $\lambda \to \infty$, the strong solution of this SDE converges a.s. to the solution of the ODE

$$
y' = \frac{1}{2}\sqrt{\cosh^2 y - a}, \quad y(0) = -\infty.
$$

This ODE is can be solved and the solutions satisfies $\int_{-\infty}^{y(t)} \frac{2}{\sqrt{\cosh^2 x - a}}dx = t$. This shows that y explodes exactly at

$$
\int_{-\infty}^{\infty} \frac{2}{\sqrt{\cosh^2 x - a}}dx = 4K(a) = t_a.
$$

This shows that $\lim_{\lambda\to\infty} P(\lambda\tau_{Y,\lambda} \in (t_a - \varepsilon, t_a + \varepsilon)) = 1$ for fixed a and ε and this completes the proof of the proposition. □

We can use the tail estimates of τ_λ to estimate the tail probabilities of $\widetilde{\alpha}_\lambda(t)$ for a fixed t. Recall the definition of $\mathcal{I}(\cdot)$ from (4.3).

Lemma 4.9. *There exist a constant c so that for $\lambda > 2$ we have*

$$e^{-\lambda^2 t\mathcal{I}(q)+\lambda c(t+1)(\mathcal{I}(q)+1)} \geq \begin{cases} P(\lceil\widetilde{\alpha}_\lambda(t)\rceil_{2\pi} \geq qt\lambda) & \text{if } q > 1, \\ P(\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi} \leq qt\lambda) & \text{if } 0 < q < 1. \end{cases} \tag{4.21}$$

Moreover, there are absolute constants c_0, c_1 so that if $qt\lambda, q$ and $\lambda q \log q$ are all bigger than c_0 then

$$P(\lceil\widetilde{\alpha}_\lambda(t)\rceil_{2\pi} \geq qt\lambda) \leq e^{-c_1\lambda^2 t\, q^2 \log q}. \tag{4.22}$$

Proof. Introduce the hitting times

$$\tau_\lambda^{(n)} = \inf\{t > 0 : \widetilde{\alpha}_\lambda(t) > 2n\pi\}. \tag{4.23}$$

Then by Proposition 4.2 the random variables $\tilde{\tau}^{(n)} = \tau^{(n)} - \tau^{(n-1)}$ are i.i.d. with the same distribution as τ_λ. Applying the exponential Markov inequality we get

$$P(\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi} \leq qt\lambda) = P\left(\sum_{i=1}^{\lceil qt\lambda/(2\pi)\rceil} \tilde{\tau}^{(i)} \geq t\right) \leq \left(Ee^{A\tau_\lambda}\right)^{\lceil qt\lambda/(2\pi)\rceil} e^{-At} \tag{4.24}$$

with any $A > 0$. Suppose first that $q < 1$ which also implies $a = a(q) = K^{-1}(\pi/(2q)) \in (0, 1)$. By choosing

$$A = \frac{\lambda^2 a}{8} - \frac{\lambda|a|}{4} \tag{4.25}$$

we have $A > 0$ if $\lambda > 2$ and from (4.10) we have $Ee^{A\tau_\lambda} \le e^{-\lambda \mathcal{H}(a)}$. Together with (4.24) this gives

$$P(\widetilde{\alpha}_\lambda(t) \le qt\lambda) \le e^{-\lambda \mathcal{H}(a)\lceil qt\lambda/(2\pi)\rceil - (\frac{\lambda^2 a}{8} + \frac{\lambda|a|}{4})t}$$
$$\le e^{-\frac{qt\lambda^2}{2\pi}\mathcal{H}(a) - (\frac{\lambda^2 a}{8}+\frac{\lambda|a|}{4})t + \lambda|\mathcal{H}(a)|} = e^{-\lambda^2 t\mathcal{I}(q) + \lambda\left(\frac{|a|t}{4} + |\mathcal{H}(a)|\right)} \tag{4.26}$$

where we used the definitions (4.3) and (2.11).

For the $q > 1$ case we use the same steps. Here $a = K^{-1}(\pi/(2q)) < 0$ and A defined in (4.25) is negative which is exactly what we need for the exponential Markov inequality. Eventually we get

$$P(\lceil\widetilde{\alpha}_\lambda(t)\rceil_{2\pi} \ge q\lambda t) \le e^{-\lambda \mathcal{H}(a)\lfloor qt\lambda/(2\pi)\rfloor - (\frac{\lambda^2 a}{8} - \frac{\lambda|a|}{4})t}$$
$$\le e^{-\frac{qt\lambda^2}{2\pi}\mathcal{H}(a) - (\frac{\lambda^2 a}{8}-\frac{\lambda|a|}{4})t + \lambda|\mathcal{H}(a)|} = e^{-\lambda^2 t\mathcal{I}(q) + \lambda\left(\frac{|a|t}{4} + |\mathcal{H}(a)|\right)}. \tag{4.27}$$

By Lemma B.3 in the Appendix there is a constant c so that

$$\mathcal{H}(a(q)) + \tfrac{1}{4}|a(q)|t \le c(t+1)(\mathcal{I}(q)+1), \tag{4.28}$$

for all $t, q > 0$ which means that we can replace the upper bounds in (4.26) and (4.27) with $e^{-\lambda^2 t\mathcal{I}(q) + \lambda c(t+1)(\mathcal{I}(q)+1)}$. This proves the first part of Lemma 4.9.

For the second part we repeat the same steps as in the $q > 1$ case, but now use

$$A = \frac{\lambda^2 a}{8} - \frac{\lambda\sqrt{|a|}}{4}.$$

This gives

$$P(\lceil\widetilde{\alpha}_\lambda(t)\rceil_{2\pi} \ge q\lambda t) \le e^{-\lambda \mathcal{H}(a)\lfloor qt\lambda/(2\pi)\rfloor - \frac{\lambda^2 a}{8}t + \frac{\lambda\sqrt{|a|}}{4}t}.$$

By Proposition B.2 of the Appendix if q is large enough then $a = K^{-1}(\pi/(2q)) > cq^2\log^2 q$ with some positive constant c. If $-a\lambda$ and $qt\lambda$ are big enough (which can be achieved by choosing c_0 big enough), we will have

$$\lfloor qt\lambda/(2\pi)\rfloor > \frac{9}{10}qt\lambda/(2\pi), \qquad -\frac{\lambda^2 a}{8} + \frac{\lambda\sqrt{|a|}}{4} < -\frac{11}{10}\cdot\frac{\lambda^2 a}{8}.$$

Then

$$\begin{aligned}
-\lambda\mathcal{H}(a)\lfloor qt\lambda/(2\pi)\rfloor - \tfrac{1}{8}\lambda^2 at + \tfrac{1}{4}\lambda\sqrt{|a|}t &< -\lambda^2 t\left(\tfrac{9}{10}\mathcal{H}(a)\tfrac{q}{2\pi} + \tfrac{11}{10}\tfrac{a}{8}\right)\\
&= -\lambda^2 t\left(-\frac{7a}{80} - \frac{9E(a)}{40K(a)} + \frac{9}{40}\right)\\
&< -c_2\lambda^2 tq^2\log^2 q.
\end{aligned}$$

with a positive constant c_2, where in the last step we again used the asymptotics given in Proposition B.2 together with (B.5). This completes the proof of (4.22). $\square$

4.3 The path deviation for the $\widetilde{\alpha}_\lambda$ process

In this section we will prove Theorem 4.6. In order to show the large deviation principle we need that

$$\begin{aligned}
\liminf_{\lambda\to\infty}\frac{1}{\lambda^2}P\left(\frac{\widetilde{\alpha}_\lambda(\cdot)}{\lambda}\in G\right) &\geq \inf_{g\in G}\mathcal{J}_{\mathrm{Sch},T}(g), &&\text{for any open set } G\subset C[0,T],\\
\limsup_{\lambda\to\infty}\frac{1}{\lambda^2}P\left(\frac{\widetilde{\alpha}_\lambda(\cdot)}{\lambda}\in K\right) &\leq \inf_{g\in K}\mathcal{J}_{\mathrm{Sch},T}(g), &&\text{for any closed set } K\subset C[0,T].
\end{aligned}$$

The fact that $\mathcal{J}_{\mathrm{Sch},T}(g)$ is a good rate function will be proved in Proposition 4.13 of Section 4.5.

We will use the fact that $\mathcal{I}(x)$ is strictly convex on $(0,\infty)$ with a global minimum at $\mathcal{I}(1)=0$, and also that there is a constant $c>0$ so that

$$c^{-1}\leq\frac{\mathcal{I}(x)}{x^2\log^2 x}\leq c, \qquad \text{for all } x>2. \tag{4.29}$$

These statements will be proved in Propositions B.1 and B.2 of the Appendix.

Proof of the large deviations upper bound in Theorem 4.6. We will follow the standard strategy for proving path level large deviations. Consider a closed subset K of $C[0,T]$. We need to bound $P(\frac{1}{\lambda}\widetilde{\alpha}_\lambda(\cdot) \in K)$. Define the δ-'fattening' of K as

$$K^\delta := \{f \in C[0,T] : \|f-g\| \le \delta \text{ for some } g \in K\}. \tag{4.30}$$

From now on $\|\cdot\|$ denotes the sup-norm on the appropriate interval.

Let π_N be the following projection of $C[0,T]$ to piecewise linear paths:

$$(\pi_N f)(iT/N) = \lfloor f(iT/N) \rfloor_{2\pi}, \qquad 0 \le i \le N \tag{4.31}$$

and $\pi_N f$ is defined linearly between these points. Then

$$P(\tilde{\alpha}_\lambda/\lambda \in K) \le P(\|\tilde{\alpha}_\lambda - \pi_N\tilde{\alpha}_\lambda\| \ge \delta\lambda) + P\big(\pi_N(\tfrac{1}{\lambda}\tilde{\alpha}_\lambda) \in K^\delta\big). \tag{4.32}$$

We will bound the two probabilities in (4.32) separately.

The first term can be rewritten as

$$P\left[\|\tilde{\alpha}_\lambda - \pi_N\tilde{\alpha}_\lambda\| \ge \delta\lambda\right] = P\Bigg(\max_k \sup_{t\in[\frac{(k-1)T}{N},\frac{kT}{N}]} |\pi_N\tilde{\alpha}_\lambda(t) - \tilde{\alpha}_\lambda(t)| \ge \delta\lambda\Bigg). \tag{4.33}$$

By Proposition 4.2 the process $\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi}$ is non-decreasing.Thus for any fixed k we have

$$\sup_{t\in[\frac{(k-1)T}{N},\frac{kT}{N}]} |\pi_N\tilde{\alpha}_\lambda(t) - \tilde{\alpha}_\lambda(t)| \le \lceil\widetilde{\alpha}_\lambda((k+1)T/N)\rceil_{2\pi} - \lfloor\widetilde{\alpha}_\lambda(kT/N)\rfloor_{2\pi}.$$

By Proposition 4.3 the term on the right is stochastically dominated by $\widetilde{\alpha}_\lambda(T/N) + 4\pi$ therefore

$$P(\|\tilde{\alpha}_\lambda - \pi_N\tilde{\alpha}_\lambda\| \ge \delta\lambda) \le NP(\widetilde{\alpha}_\lambda(T/N) + 4\pi \ge \delta\lambda) \le NP\left(\frac{1}{\lambda}\widetilde{\alpha}_\lambda(T/N) \ge \frac{\delta}{2}\right) \tag{4.34}$$

where the last bound holds if $\lambda > 8\pi/\delta$. Using Lemma 4.9 we get

$$NP\left(\frac{1}{\lambda}\widetilde{\alpha}_\lambda(T/N) \ge \frac{\delta}{2}\right) \le Ne^{-(\lambda^2\frac{T}{N}+\lambda c_1(T/N+1))\mathcal{I}\left(\frac{\delta N}{2T}\right)+\lambda c_1(T/N-1)}$$

and this leads to

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(\|\tilde{\alpha}_\lambda - \pi_N \tilde{\alpha}_\lambda\| \ge \delta\lambda) \le -\frac{T}{N}\mathcal{I}\left(\frac{\delta N}{2T}\right). \tag{4.35}$$

Note, that for fixed δ and T as $N \to \infty$ the right hand side converges to $-\infty$ by (4.29).

The second term on the right side of (4.32) can be bounded as

$$P\left(\pi_N(\widetilde{\alpha}_\lambda/\lambda) \in K^\delta\right) \le P\left(\mathcal{J}_{\mathrm{Sch},T}(\pi_N(\widetilde{\alpha}_\lambda/\lambda)) \ge \inf_{g\in K^\delta} \mathcal{J}_{\mathrm{Sch},T}(g)\right).$$

We introduce

$$\Delta\widetilde{\alpha}_i = \frac{N}{\lambda T}\left(\lfloor \widetilde{\alpha}(iT/N)\rfloor_{2\pi} - \lfloor \widetilde{\alpha}((i-1)T/N)\rfloor_{2\pi}\right), \qquad \text{for } 1 \le i \le N$$

and $C_\delta = \inf_{g\in K^\delta} \mathcal{J}_{\mathrm{Sch},T}(g)$. Then we have to bound

$$P(\mathcal{J}_{\mathrm{Sch},T}(\pi_N(\widetilde{\alpha}_\lambda/\lambda)) \ge C_\delta) = P\left(\sum_{i=1}^N \frac{T}{N}\mathcal{I}\left(\Delta\widetilde{\alpha}_i\right) \ge C_\delta\right). \tag{4.36}$$

We can apply Proposition 4.3 with $t_i = \frac{iT}{N}, 1 \le i \le N$ to get independent random variables ξ_i with $\xi_i \stackrel{d}{=} \widetilde{\alpha}_\lambda(T/N)$ and

$$\frac{N}{\lambda T}\lfloor \xi_i \rfloor_{2\pi} \le \Delta\widetilde{\alpha}_i \le \frac{N}{\lambda T}\left(\lfloor \xi_i\rfloor_{2\pi} + 2\pi\right).$$

Because of the convexity of $\mathcal{I}(\cdot)$ we then have

$$\begin{aligned}\mathcal{I}\left(\Delta\widetilde{\alpha}_i\right) &\le \max\left(\mathcal{I}\left(\frac{N\lfloor \xi_i\rfloor_{2\pi}}{\lambda T}\right), \mathcal{I}\left(\frac{N\lfloor \xi_i\rfloor_{2\pi}}{\lambda T} + \frac{2\pi N}{\lambda T}\right)\right)\\ &\le (1 + \frac{2\pi N}{\lambda T})\mathcal{I}\left(\frac{N\lfloor \xi_i\rfloor_{2\pi}}{\lambda T}\right) + c\frac{2\pi N}{\lambda T}\end{aligned}$$

where we used Lemma B.4 of the Appendix for the last bound. Fix $1/2 > \varepsilon > 0$. Using the exponential Markov inequality, the independence of ξ_i and $\xi_i \stackrel{d}{=} \widetilde{\alpha}_\lambda(T/N)$ we get the bound

$$P\left(\sum_{i=1}^{N}\frac{T}{N}\mathcal{I}\left(\Delta\widetilde{\alpha}_i\right)\geq C_\delta\right)\leq\left(Ee^{(1-2\varepsilon)\lambda^2\frac{T}{N}\left((1+\frac{2\pi N}{\lambda T})\mathcal{I}\left(\frac{N\lfloor\widetilde{\alpha}_\lambda(T/N)\rfloor_{2\pi}}{\lambda T}\right)+c\frac{2\pi N}{\lambda T}\right)}\right)^N e^{-(1-2\varepsilon)\lambda^2C_\delta}$$
$$\leq\left(Ee^{(1-\varepsilon)\lambda^2\frac{T}{N}\mathcal{I}\left(\frac{N\lfloor\widetilde{\alpha}_\lambda(T/N)\rfloor_{2\pi}}{\lambda T}\right)}\right)^N e^{(1-2\varepsilon)c2\pi\lambda N-(1-2\varepsilon)\lambda^2C_\delta},\tag{4.37}$$

where the second inequality holds for fixed ε, N, T if λ is big enough. Our next step is to estimate the exponential moment $Ee^{(1-\varepsilon)\lambda^2\frac{T}{N}\mathcal{I}\left(\frac{N\lfloor\widetilde{\alpha}_\lambda(T/N)\rfloor_{2\pi}}{\lambda T}\right)}$ for a fixed $\varepsilon>0$. By Lemma 4.10 below if N, T, ε are fixed then

$$\limsup_{\lambda\to\infty}\frac{1}{\lambda^2}\log Ee^{(1-\varepsilon)\lambda^2\frac{T}{N}\mathcal{I}\left(\frac{N\lfloor\widetilde{\alpha}_\lambda(T/N)\rfloor_{2\pi}}{\lambda T}\right)}\leq 0.$$

Using this with (4.37) we get

$$\limsup_{\lambda\to\infty}\frac{1}{\lambda^2}\log P\left(\sum_{i=1}^{N}\frac{T}{N}\mathcal{I}\left(\Delta\widetilde{\alpha}_i\right)\geq C_\delta\right)\leq -(1-2\varepsilon)C_\delta.\tag{4.38}$$

Now we let $\varepsilon\to 0$ and then $N\to\infty$. The bounds (4.35), (4.38) with (4.32) give

$$\limsup_{\lambda\to\infty}\frac{1}{\lambda^2}\log P(\lambda^{-1}\widetilde{\alpha}_\lambda(\cdot)\in K)\leq -\inf_{g\in K^\delta}\mathcal{J}_{\mathrm{Sch},T}(g).\tag{4.39}$$

Using the fact that $\mathcal{J}_{\mathrm{Sch},T}$ is a good rate function (which is proved in Proposition 4.13 of Section 4.5) we get that the right hand side converges to $-\inf_{g\in K}\mathcal{J}_{\mathrm{Sch},T}(g)$ as $\delta\to 0$. (See e.g. Lemma 4.1.6 from [14].) This finishes the proof of the lower bound. □

Now we will prove the missing estimate for the lower bound.

Lemma 4.10. *Fix $t>0$ and $1>\varepsilon>0$. Then*

$$\limsup_{\lambda\to\infty}\frac{1}{\lambda^2}\log Ee^{(1-\varepsilon)\lambda^2 t\mathcal{I}\left(\frac{\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi}}{\lambda t}\right)}\leq 0.$$

Proof. Introduce the temporary notation $G(x) = \lambda^2 t\mathcal{I}(x)$. This is a convex function with $G(1) = 0$ as its minimum. Then we have

$$Ee^{(1-\varepsilon)\lambda^2 t\mathcal{I}\left(\frac{\lfloor\widetilde{\alpha}_\lambda(t)}{\lambda t}\right)\rfloor_{2\pi}} \leq 2 - \int_0^1 (1-\varepsilon)G'(x)e^{(1-\varepsilon)G(x)}P(\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi} < \lambda tx)dx$$
$$+ \int_1^\infty (1-\varepsilon)G'(x)e^{(1-\varepsilon)G(x)}P(\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi} > \lambda tx)dx.$$

Using Lemma 4.9 we get

$$P(\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi} < \lambda tx) \leq \exp\left\{-(1-c_1\lambda^{-1}(1+t^{-1}))G(x) + \lambda c_1(t+1)\right\}$$

for $x < 1$ and a similar bound for $P(\lfloor\widetilde{\alpha}_\lambda(t)\rfloor_{2\pi} > \lambda tx) \leq P(\lceil\widetilde{\alpha}_\lambda(t)\rceil_{2\pi} > \lambda tx)$ for $x > 1$. This gives us

$$Ee^{(1-\varepsilon)\lambda^2 t\mathcal{I}\left(\frac{\widetilde{\alpha}_\lambda(t)}{\lambda t}\right)} \leq 2 - \int_0^1 (1-\varepsilon)G'(x)e^{((1+t^{-1})(c_1/\lambda)-\varepsilon)G(x)+\lambda c_1(t+1)}dx$$
$$+ \int_1^\infty (1-\varepsilon)G'(x)e^{((1+t^{-1})(c_1/\lambda)-\varepsilon)G(x)+\lambda c_1(t+1)}dx$$
$$\leq 2 + 4\varepsilon^{-1}e^{\lambda c_1(t+1)}$$

where the last inequality holds if $(1+t^{-1})c_1/\lambda < \varepsilon/2$, i.e. for large enough λ. From this the lemma follows. $\square$

Now we turn to the lower bound proof in the large deviation result of Theorem 4.6. As we will see, we will be able to reduce the problem to studying the probability of $\frac{1}{\lambda}\widetilde{\alpha}_\lambda(t)$ being close to a straight line.

Proposition 4.11. *Fix $q > 0$ and $T, \varepsilon > 0$. Then*

$$\lim_{\varepsilon\to 0}\liminf_{\lambda\to\infty} \frac{1}{\lambda^2}\log P(\widetilde{\alpha}(t) \in [\lambda(qt-\varepsilon), \lambda(qt+\varepsilon)], t \in [0,T]) \geq -T\mathcal{I}(q).$$

Proof. As $q > 0$ we may assume $\varepsilon \le qT/2$ by choosing ε small enough. Let $N = \frac{\lceil (qT+\epsilon)\lambda \rceil_{2\pi}}{2\pi}$, and choose $\varepsilon_1 = \frac{\pi\varepsilon}{2q(qT+\varepsilon)}$, which satisfies $\varepsilon_1 < \frac{\varepsilon\lambda}{2qN}$ for $\lambda > 2$. Recall the definition of $\tau_\lambda^{(n)}$ and $\widetilde{\tau}_\lambda^{(n)}$ from (4.23). We will prove that

$$\begin{aligned} P(\widetilde{\alpha}(t) \in [\lambda(qt-\varepsilon), &\lambda(qt+\varepsilon)], t \in [0,T]) \\ &\ge P\left(\lambda\widetilde{\tau}_\lambda^{(k)} \in (\tfrac{2\pi}{q} - \varepsilon_1, \tfrac{2\pi}{q} + \varepsilon_1), 1 \le k \le N\right). \end{aligned}$$

Roughly speaking, this will follow from the simple fact that if we are within ε/q of the line $y = qt$ in the horizontal direction, then we are within ε in the vertical direction. If $\lambda\widetilde{\tau}_\lambda^{(k)} \ge \frac{2\pi}{q} - \varepsilon_1$ for $1 \le k \le N$ then $\lambda\tau_\lambda^{(k)} \ge k(\frac{2\pi}{q} - \varepsilon_1)$ and

$$\widetilde{\alpha}_\lambda\left(\tfrac{k}{\lambda}\left(\tfrac{2\pi}{q} - \varepsilon_1\right)\right) \le 2k\pi = \lambda\frac{2\pi}{2\pi/q - \varepsilon_1} \cdot \frac{k}{\lambda}\left(\tfrac{2\pi}{q} - \varepsilon_1\right)$$

for $1 \le k \le N$. Together with the fact that $\lfloor \widetilde{\alpha}_\lambda \rfloor_{2\pi}$ is non-decreasing we get that

$$\widetilde{\alpha}_\lambda(t) \le \lambda\frac{2\pi}{2\pi/q - \varepsilon_1} \cdot \left(t + \tfrac{1}{\lambda}(2\pi/q - \varepsilon_1)\right), \qquad \text{for } t \le \tfrac{N}{\lambda}\left(\tfrac{2\pi}{q} - \varepsilon_1\right).$$

This inequality implies $\widetilde{\alpha}_\lambda(t) \le \lambda(qt+\varepsilon)$, for $t \le T$, $\lambda\varepsilon > 4\pi$.

The other direction is similar, if we have $\lambda\widetilde{\tau}_\lambda^{(k)} \le \frac{2\pi}{q} + \varepsilon_1$ for $1 \le k \le N$ then

$$\widetilde{\alpha}_\lambda\left(\tfrac{k}{\lambda}\left(\tfrac{2\pi}{q} + \varepsilon_1\right)\right) \ge 2k\pi = \lambda\frac{2\pi}{2\pi/q + \varepsilon_1} \cdot \frac{k}{\lambda}\left(\tfrac{2\pi}{q} + \varepsilon_1\right)$$

which implies

$$\widetilde{\alpha}_\lambda(t) \ge \lambda\frac{2\pi}{2\pi/q + \varepsilon_1} \cdot \left(t - \tfrac{1}{\lambda}(2\pi/q + \varepsilon_1)\right), \qquad \text{for } t \le \tfrac{N}{\lambda}\left(\tfrac{2\pi}{q} + \varepsilon_1\right).$$

and $\widetilde{\alpha}_\lambda(t) \ge \lambda(qt-\varepsilon)$ for $t \le T$. Using the independence of $\widetilde{\tau}_\lambda^{(k)}$ we get the bound

$$P(\widetilde{\alpha}(t) \in [\lambda(qt-\varepsilon), \lambda(qt+\varepsilon)], t \in [0,T]) \ge P\left(\lambda\widetilde{\tau}_\lambda^{(k)} \in (\tfrac{2\pi}{q} - \varepsilon_1, \tfrac{2\pi}{q} + \varepsilon_1)\right)^N. \quad (4.40)$$

By the lower bound (4.11) we have

$$\log P(\widetilde{\alpha}(t) \in [\lambda(qt-\varepsilon), \lambda(qt+\varepsilon)], t \in [0,T]) \\ \geq \frac{\lambda(qT+2\varepsilon)}{2\pi}\left(-\frac{2\pi\lambda}{q}\mathcal{I}(q) - \frac{\lambda|a|}{8}(\varepsilon_1 + 4(2\pi/q+\varepsilon_1)) + \log A(\varepsilon_1, \lambda, a)\right).$$

Recalling $\varepsilon_1 = \frac{\pi\varepsilon}{2q(qT+\varepsilon)}$ we get

$$\lim_{\varepsilon\to 0}\liminf_{\lambda\to\infty} \frac{1}{\lambda^2}\log P(\widetilde{\alpha}(t) \in [\lambda(qt-\varepsilon), \lambda(qt+\varepsilon)], t \in [0,T]) \geq -T\mathcal{I}(q). \qquad \square$$

Proof of the lower bound in Theorem 4.6. Let G be an open subset of $C[0,T]$. We would like to show that

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda^2}\log P(\frac{1}{\lambda}\widetilde{\alpha}_\lambda(\cdot) \in G) \geq -\inf_{g\in G}\int_0^T \mathcal{I}\left(g'(t)\right)dt. \tag{4.41}$$

For this it is enough to prove that for any $g \in G$ with $\int_0^T \mathcal{I}\left(g'(t)\right)dt < \infty$ and $\delta > 0$ we have

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda^2}\log P(\frac{1}{\lambda}\widetilde{\alpha}_\lambda(\cdot) \in G) \geq -\int_0^T \mathcal{I}\left(g'(t)\right)dt - \delta. \tag{4.42}$$

We can approximate g with a piecewise linear function $\tilde{g}$ in the sup-norm so that we have $|\int_0^T \mathcal{I}\left(g_n'(t)\right)dt - \int_0^T \mathcal{I}\left(g'(t)\right)dt| < \delta$. Because of this we may assume that g is piecewise linear, moreover, we may assume that there are no horizontal segments in g. Suppose that g is linear with slope q_i on the interval $[T_i, T_{i+1}]$ with $0 \leq i \leq k-1$ and $0 = T_0 < T_1 < \cdots < T_k = T$. We claim that if $\lambda > \lambda_0(\varepsilon, k)$ then

$$P(\|\frac{1}{\lambda}\widetilde{\alpha}_\lambda(\cdot) - g(\cdot)\| \leq \varepsilon) \geq P\left(|\frac{1}{\lambda}(\widetilde{\alpha}_\lambda(t) - \widetilde{\alpha}_\lambda(T_i)) - q_i(t-T_i)| \leq \varepsilon/k, \text{ if } t \in [T_i, T_{i+1}]\right) \\ \geq \prod_{i=0}^{k-1} P\left(|\frac{1}{\lambda}\widetilde{\alpha}_\lambda(t) - q_i t| \leq \varepsilon/(2k), \text{ for } t \in [0, T_{i+1}-T_i]\right) \tag{4.43}$$

The first inequality is straightforward, to prove the second we use the coupling in the proof of Proposition 4.3. Recall the definition of the processes $\hat{\alpha}_i(s)$ defined on $[t_{i-1}, t_i]$. These were independent for different values of i and the process $\hat{\alpha}_i(s+t_{i-1}), s \in [0, t_i - t_{i-1}]$ had the same distribution as $\widetilde{\alpha}_\lambda(s), s \in [0, t_i - t_{i-1}]$. We also had

$$\hat{\alpha}_i(s) + \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi} \leq \widetilde{\alpha}_\lambda(s) \leq \hat{\alpha}_i(s) + \lfloor \widetilde{\alpha}_\lambda(t_{i-1}) \rfloor_{2\pi} + 2\pi.$$

for $s \in [t_{i-1}, t_i]$. By choosing $\lambda > \lambda_0 = 4\pi k/\varepsilon$ the inequality (4.43) follows by the independent increment property of the Brownian motion.

By Proposition 4.11 we have the bound

$$\lim_{\varepsilon \to 0} \liminf_{\lambda \to \infty} \tfrac{1}{\lambda^2} \log P(\| \frac{1}{\lambda} \widetilde{\alpha}_\lambda(\cdot) - g(\cdot) \| \leq \varepsilon) \geq - \sum_{i=0}^{k-1} (T_{i+1} - T_i) \mathcal{I}(q_i) = - \int_0^T \mathcal{I}\left(g'(t)\right) dt$$

from which (4.42) and thus the proof of the lower bound follows. □

4.4 The path deviation for the Sine_β process

This section contains the proof of Theorem 4.5. The strategy for the proof is to approximate the SDE (2.9) with a version where the drift is piecewise constant and then use elements of the proof of Theorem 4.6. Just as in the proof of Theorem 4.6, we need to show an upper and a lower bound to prove the large deviation principle. The fact that $\mathcal{J}_{\mathrm{Sine}_\beta}$ is a good rate function will be proved in Proposition 4.13 of Section 4.5.

Proof of the upper bound in Theorem 4.5. For the proof of the upper bound we go through a series of approximations: we essentially cut of the tail of the process, then replace the drift in the SDE with a piecewise constant version and then approximate the process with a piecewise linear version. Recall that $\alpha_\lambda(t)$ solves the SDE (2.8) and that we introduced the notation $\mathfrak{f}(t) = \frac{\beta}{4}e^{\frac{\beta}{4}t}$. Fix $T > 0$, the value of which will go to infinity later. The first approximating process is defined as

$$\alpha_\lambda^{(1)}(t) = \alpha_\lambda(t)\mathbf{1}(t \le T) + (\alpha_\lambda(T) + \lambda(e^{-\frac{\beta}{4}T} - e^{-\frac{\beta}{4}t}))\mathbf{1}(t > T),$$

this solves the SDE (2.8) with the noise 'turned off' at $t = T$. For the second process we define

$$\mathfrak{f}_N(t) = \mathfrak{f}(Ti/N), \qquad t \in [Ti/N, T(i+1)/N) \tag{4.44}$$

and consider the solution $\xi_{\lambda\mathfrak{f}_N}$ of (4.7) with drift $\lambda\mathfrak{f}_N$ and initial condition 0. Let

$$\alpha_\lambda^{(2)}(t) = \xi_{\lambda\mathfrak{f}_N}(t)\mathbf{1}(t \le T) + (\xi_{\lambda\mathfrak{f}_N}(T) + \lambda(e^{-\frac{\beta}{4}T} - e^{-\frac{\beta}{4}t}))\mathbf{1}(t > T).$$

Finally, let π_{MN} is the projection defined in (4.31) with intervals of size T/MN, that is $\pi_{MN}f$ is the piecewise linear path that satisfies

$$(\pi_{MN}f)(Ti/(MN)) = \lfloor f(Ti/(MN)) \rfloor_{2\pi},$$

and is linear between these values. Define

$$\alpha_\lambda^{(3)}(t) = \pi_{MN}\xi_{\lambda\mathfrak{f}_N}(t)\mathbf{1}(t \le T) + (\pi_{MN}\xi_{\lambda\mathfrak{f}_N}(T) + \lambda(e^{-\frac{\beta}{4}T} - e^{-\frac{\beta}{4}t}))\mathbf{1}(t > T).$$

Then for any closed set $K \subset C[0,\infty)$ we have that

$$\begin{aligned} P\left(\frac{\alpha_\lambda}{\lambda} \in K\right) \le P\Big(\frac{\alpha_\lambda^{(3)}}{\lambda} \in K^{3\delta}\Big) &+ P(\|\alpha_\lambda^{(1)} - \alpha_\lambda\|_\infty \ge \delta\lambda) \\ &+ P(\|\alpha_\lambda^{(2)} - \alpha_\lambda^{(1)}\|_\infty \ge \delta\lambda) + P(\|\alpha_\lambda^{(3)} - \alpha_\lambda^{(2)}\|_\infty \ge \delta\lambda), \end{aligned} \tag{4.45}$$

where $K^{3\delta}$ is defined similarly to (4.30), as the 3δ-fattening of K. We will begin with the main term. Let

$$\mathcal{J}_N(g) = \int_0^\infty \mathfrak{f}_N^2(t)\mathcal{I}\left(\frac{g'(t)}{\mathfrak{f}_N(t)}\right) dt,$$

and define (similarly to the $\tilde{\alpha}_\lambda$ case in the proof of Theorem 4.6)

$$\Delta\alpha_i = \frac{MN}{\lambda\mathfrak{f}_N(\frac{Ti}{MN})T}\left(\lfloor\alpha^{(3)}(Ti/(MN))\rfloor_{2\pi} - \lfloor\alpha^{(3)}(T(i-1)/(MN))\rfloor_{2\pi}\right), \text{ for } 1 \le i \le MN.$$

Then,

$$\begin{aligned} P\Big(\frac{\alpha_\lambda^{(3)}}{\lambda} \in K^{3\delta}\Big) &\le P\left(\mathcal{J}_N\Big(\frac{\alpha_\lambda^{(3)}}{\lambda}\Big) \ge \inf_{g\in K^{3\delta}} \mathcal{J}_N(g)\right) \\ &= P\left(\sum_{i=1}^{MN} \frac{T(\mathfrak{f}_N(Ti/(MN)))^2}{MN}\mathcal{I}(\Delta\alpha_i) \ge \inf_{g\in K^{3\delta}} \mathcal{J}_N(g)\right). \end{aligned}$$

Take $\widehat{\alpha}_i$ to solve (4.4) but with the Brownian motion $B(t+Ti/(MN)) - B(Ti/(MN))$ and $\lambda_i = \lambda\mathfrak{f}_N(Ti/(MN))$. Then using the same arguments as in the bound (4.37) we get

$$P\Big(\frac{\alpha_\lambda^{(3)}}{\lambda} \in K^{3\delta}\Big) \le e^{-(1-\varepsilon)\lambda^2 C_{\delta,N}} \prod_{i=1}^{MN}\left(Ee^{(1-\varepsilon)\lambda_i^2\frac{T}{MN}\left((1+\frac{2\pi MN}{\lambda T})\mathcal{I}\left(\frac{\lfloor\widehat{\alpha}_i(T/(MN))\rfloor_{2\pi}}{\lambda_i(T/MN)}\right)+c_2\frac{2\pi MN}{\lambda_i T}\right)}\right)$$

where $C_{\delta,N} = \inf_{g\in K^{3\delta}} \mathcal{J}_N(g)$. Using the bound proved in Lemma 4.10 we get that

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2}\log P\Big(\frac{\alpha_\lambda^{(3)}}{\lambda} \in K^{3\delta}\Big) \le -(1-\varepsilon)C_{\delta,N} \tag{4.46}$$

We now turn to the first error term. Using the fact that $\lfloor\alpha_\lambda\rfloor_{2\pi}$ is non-decreasing (which follows from (i) of Proposition 4.4) we get that

$$\|\alpha_\lambda^{(1)} - \alpha_\lambda\| \le \alpha_\lambda(\infty) - \alpha_\lambda(T) + \lambda e^{-\frac{\beta}{4}T},$$

where $\alpha_\lambda(\infty)$ is the limit of $\alpha_\lambda(t)$ as $t\to\infty$. Choose T large enough so that $e^{-\frac{\beta}{4}T} \le \delta/2$. Then

$$P(\|\alpha_\lambda^{(1)} - \alpha_\lambda\| \ge \delta\lambda) \le P(\alpha_\lambda(\infty) - \alpha_\lambda(T) \ge \delta\lambda/2)$$

We will deal with this tail probability in Proposition 4.12 below. In particular, we will show that there is a constant $c_1 > 0$ so that

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(\|\alpha_\lambda^{(1)} - \alpha_\lambda\| \geq \delta\lambda) \leq -c_1 T\delta^2. \tag{4.47}$$

For the second error term we first note that $\|\alpha_\lambda^{(2)} - \alpha_\lambda^{(1)}\| = \sup_{t\in[0,T]} |\alpha_\lambda^{(2)}(t) - \alpha_\lambda^{(1)}(t)|$. Using the coupling of Proposition 4.4 we can show that on $[0,T]$ the process $\alpha_\lambda^{(2)} - \alpha_\lambda^{(1)}$ will have the same distribution as the solution of the SDE (4.7) with initial condition 0 and drift $\lambda(\mathfrak{f}_N - \mathfrak{f}) \geq 0$. Moreover, this process will be non-negative (because the drift is non-negative), and since $\lambda(\mathfrak{f}_N(t) - \mathfrak{f}(t)) \leq \lambda\frac{\beta T}{4N}$ for $t \in [0,T]$, it will be bounded by the solution of the SDE (4.7) with a constant drift $\lambda\frac{\beta T}{4N}$. Because of this $\|\alpha_\lambda^{(2)} - \alpha_\lambda^{(1)}\|$ is stochastically bounded by $\sup_{t\in[0,T]} \widetilde{\alpha}_{\lambda\frac{\beta T}{4N}}(t) \leq \widetilde{\alpha}_{\lambda\frac{\beta T}{4N}}(T) + 2\pi$ with $\widetilde{\alpha}_\lambda$ from (4.4, using the fact that $\lfloor \widetilde{\alpha}_\lambda(t) \rfloor_{2\pi}$ is non-decreasing. Thus for $\delta\lambda > 4\pi$ we have

$$P(\|\alpha_\lambda^{(2)} - \alpha_\lambda^{(1)}\| \geq \delta\lambda) \leq P(\tilde{\alpha}_{\lambda\frac{\beta T}{4N}}(T) \geq \tfrac{1}{2}\delta\lambda).$$

If N and T are fixed then if λ is big enough then we can apply Lemma 4.9 for the right hand side with $\tilde{\lambda} = \lambda\frac{\beta T}{4N}$, $t = T$ and $q = \frac{\frac{1}{2}\delta\lambda}{T\lambda\frac{\beta T}{4N}} = \frac{2\delta N}{\beta T^2}$. This leads to

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(\|\alpha_\lambda^{(2)} - \alpha_\lambda^{(1)}\| \geq \tfrac{1}{2}\delta\lambda) \leq -\frac{\beta^2 T^3}{4^2 N^2}\mathcal{I}\left(\frac{2\delta N}{\beta T^2}\right). \tag{4.48}$$

For the third error term we first note that

$$\|\alpha_\lambda^{(3)} - \alpha_\lambda^{(2)}\| \leq \sup_{t\in[0,T]} |\alpha_\lambda^{(3)}(t) - \alpha_\lambda^{(2)}(t)| \leq \max_i \sup_{t\in[Ti/N, T(i+1)/N]} |\alpha_\lambda^{(3)}(t) - \alpha_\lambda^{(2)}(t)|,$$

and thus

$$P(\|\alpha_\lambda^{(3)} - \alpha_\lambda^{(2)}\| \geq \delta\lambda) \leq \sum_{i=0}^{N-1} P\left(\sup_{t\in[Ti/N, T(i+1)/N]} |\alpha_\lambda^{(3)}(t) - \alpha_\lambda^{(2)}(t)| \geq \delta\lambda\right).$$

In the interval $[Ti/N, T(i+1)/N]$ the process $\alpha_\lambda^{(2)}$ solves the SDE (4.4) with constant drift $\lambda\mathfrak{f}_N(Ti/N)$. Here we can use the same steps that we used in the proof of Theorem 4.6 between (4.33) and (4.34) to get

$$\begin{aligned} P(\|\alpha_\lambda^{(3)} - \alpha_\lambda^{(2)}\| \geq \delta\lambda) &\leq \sum_{i=1}^{N-1} MP\left(\tilde{\alpha}_{\lambda\mathfrak{f}_N(Ti/N)}(T/(MN)) \geq \delta\lambda/2\right) \\ &\leq MNP\left(\tilde{\alpha}_{\frac{\beta}{4}\lambda}(T/(MN)) \geq \delta\lambda/2\right) \end{aligned}$$

for λ big enough compared to δ^{-1}. For large enough λ we can apply Lemma 4.9 for the right hand side with $\tilde{\lambda} = \frac{\beta}{4}\lambda$, $t = T/(MN)$ and $q = \frac{2\delta MN}{\beta T}$ to get

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(\|\alpha_\lambda^{(3)} - \alpha_\lambda^{(2)}\| \geq \delta\lambda) \leq -\frac{\beta^2}{4^2}\frac{T}{MN}\mathcal{I}\left(\frac{2\delta MN}{\beta T}\right). \tag{4.49}$$

Now taking (4.45) with the bounds (4.46), (4.47), (4.48) and (4.49) we get

$$\begin{aligned} &\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P\Big(\frac{\alpha_\lambda}{\lambda} \in K\Big) \\ &\quad\leq \max\left\{-(1-\varepsilon)C_{\delta,N}, -c_1 T\delta^2, -\frac{\beta^2 T^3}{4^2 N^2}\mathcal{I}\left(\frac{2\delta N}{\beta T^2}\right), -\frac{\beta^2}{4^2}\frac{T}{MN}\mathcal{I}\left(\frac{2\delta MN}{\beta T}\right)\right\}. \end{aligned} \tag{4.50}$$

Taking N to ∞ the last two terms go to $-\infty$ (using the bounds (4.29)) while the first term converges to $(1-\varepsilon)C_\delta^T$ with

$$C_\delta^T = \inf_{g\in K^{3\delta}} \int_0^T \mathfrak{f}^2(t)\mathcal{I}\left(g'(t)/\mathfrak{f}(t)\right) dt$$

Letting now $T \to \infty$ and then $\varepsilon \to 0$ we get

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P\Big(\frac{\alpha_\lambda}{\lambda} \in K\Big) \leq -\inf_{g\in K^{3\delta}} \mathcal{J}_{\mathrm{Sine}_\beta}(g).$$

Finally taking $\delta \to 0$ and using the fact that $\mathcal{J}_{\mathrm{Sine}_\beta}$ is a good rate function gives the result

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P\Big(\frac{\alpha_\lambda}{\lambda} \in K\Big) \leq -\inf_{g\in K} \mathcal{J}_{\mathrm{Sine}_\beta}(g).$$

This completes the proof of the lower bound. □

We now prove the tail bound for the proof of the lower bound.

Proposition 4.12. *Fix $T, \delta > 0$, then there is a constant $c > 0$ so that*

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(\alpha_\lambda(\infty) - \alpha_\lambda(T) \geq \delta\lambda) \leq -cT\delta^2. \tag{4.51}$$

Proof of Proposition 4.12. Take $\nu = 1/8$, and set $T_k = \frac{k(k+1)}{2}\theta T$ where the value of $\theta > 0$ will be specified later. Then we can break up the probability in question as

$$P(\alpha_\lambda(\infty) - \alpha_\lambda(T) \geq \delta\lambda) \leq \sum_{k=1}^{\lfloor 2\sqrt{\lambda}\rfloor} P(\alpha_\lambda(T_{k+1}) - \alpha_\lambda(T_k) \geq \tfrac{\delta}{4}\lambda\nu k^{-(1+\nu)})$$
$$+ P(\alpha_\lambda(\infty) - \alpha_\lambda(T_{\lfloor 2\sqrt{\lambda}\rfloor+1}) \geq \delta\lambda/2). \tag{4.52}$$

Note, that for any fixed $s > 0$ the process $\widehat{\alpha}_{s,\lambda}(t) = \alpha_\lambda(s+t)$ satisfies the SDE (2.9) with $\hat{\lambda} = \lambda e^{-\frac{\beta}{4}s}$ with initial condition $\alpha_\lambda(s)$. Using the coupling techniques of Propositions 4.3 and 4.4 one can show that $\widehat{\alpha}_{s,\lambda}(t) - \widehat{\alpha}_{s,\lambda}(0) = \alpha_\lambda(s+t) - \alpha_\lambda(s)$ is stochastically dominated by $\widetilde{\alpha}_{\lambda\mathfrak{f}(s)} + 2\pi$. This (together with $T_{k+1} - T_k = \theta(k+1)T$) gives

$$P(\alpha_\lambda(T_{k+1}) - \alpha_\lambda(T_k) \geq \tfrac{\delta}{4}\lambda\nu k^{-(1+\nu)}) \leq P(\widetilde{\alpha}_{\lambda\mathfrak{f}(T_k)}(\theta(k+1)T) \geq \tfrac{\delta}{4}\lambda\nu k^{-(1+\nu)} - 2\pi)$$
$$\leq P(\widetilde{\alpha}_{\lambda\mathfrak{f}(T_k)}(\theta(k+1)T) \geq \tfrac{\delta}{8}\lambda\nu k^{-(1+\nu)})$$

where the last bound follows for big enough λ from $k \leq 2\sqrt{\lambda}$. We can use bound (4.22) of Lemma 4.9 for the probability on the right with $\tilde{\lambda} = \lambda\mathfrak{f}(T_k)$, $t = \theta(k+1)T$ and $q = \frac{\delta\nu k^{-(1+\nu)}}{8\theta T(k+1)\mathfrak{f}(T_k)}$, since with these choices $qt\tilde{\lambda}$, q and $\tilde{\lambda}q\log q$ are all big, if we choose $\theta > 0$ small enough and then λ big enough. This leads to

$$P(\widetilde{\alpha}_{\lambda\mathfrak{f}(T_k)}(T) \geq \tfrac{\delta}{4}\lambda\nu k^{-(1+\nu)})$$
$$\leq \exp\left(-c_1\tfrac{\delta^2}{8^2}\lambda^2\nu^2k^{-2(1+\nu)}\theta^{-1}(k+1)^{-1}T^{-1}\log^2\left(\frac{\delta\nu k^{-(1+\nu)}}{8\theta(k+1)T\mathfrak{f}(T_k)}\right)\right)$$
$$\leq \exp\left(-c_2\delta^2\lambda^2k^{-3-2\nu}T^{-1}\left(c_3 + \tfrac{\beta}{4}T\tfrac{k(k+1)}{2}\right)^2\right) \leq \exp\left(-c_4\lambda^2\delta^2Tk^{1-2\nu}\right),$$

with a positive constant c_4, which in turn implies (for large enough λ)

$$\sum_{k=1}^{\lfloor 2\sqrt{\lambda}\rfloor} P(\alpha_\lambda(T_{k+1}) - \alpha_\lambda(T_k) \geq \tfrac{\delta}{4}\lambda\nu k^{-(1+\nu)}) \leq 2\exp\left(-c_4\lambda^2\delta^2 T\right). \tag{4.53}$$

Lastly we bound the remaining term using Proposition 4.4:

$$P(\alpha(\infty) - \alpha(T\lambda) \geq \delta\lambda/2) = P(\xi_{\lambda\mathfrak{f}(\lambda T)}(\infty) \geq \lfloor \delta\lambda/2 \rfloor) \leq 2\left(e^{-\frac{\beta}{4}\lambda T}\right)^{\lfloor \delta\lambda/2\rfloor},$$

which together with (4.52) and (4.53) gives us the necessary upper bound for (4.51). □

Proof of the lower bound in Theorem 4.5. We will show that if $g \in C[0,\infty)$ with $\mathcal{J}_{\mathrm{Sine}_\beta}(g) < \infty$ then

$$\lim_{\varepsilon\to 0}\liminf_{\lambda\to\infty} \tfrac{1}{\lambda^2}\log P(\|\lambda^{-1}\alpha_\lambda(\cdot) - g(\cdot)\| \leq \varepsilon) \geq -\mathcal{J}_{\mathrm{Sine}_\beta}(g). \tag{4.54}$$

From this the lower bound will follow.

In Proposition 4.13 of the Appendix we will prove that if $\mathcal{J}_{\mathrm{Sine}_\beta}(g) < \infty$ then $g(\infty) = \lim_{t\to\infty} g(t) < \infty$ exists. Let $\varepsilon > 0$ and choose $T > 0$ so that

$$g(\infty) - g(T) \leq \varepsilon/2, \qquad \text{and} \qquad e^{-\frac{\beta}{4}T} \leq \varepsilon/4. \tag{4.55}$$

From the first assumption in (4.55) and the Markov property we have

$$\begin{aligned}
&P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \leq \varepsilon, t \geq 0)\\
&\quad \geq P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \leq \varepsilon/2, t \in [0,T], |\alpha_\lambda(\infty) - \alpha_\lambda(T)| \leq \lambda\varepsilon/4) \qquad (4.56)\\
&\quad \geq P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \leq \varepsilon/2, t \in [0,T]) \sup_x P(\alpha_\lambda(\infty) - \alpha_\lambda(T) \leq \lambda\varepsilon/4 \big| \alpha_\lambda(T) = x).
\end{aligned}$$

Using the same line of reasoning as in the proof of Proposition 4.12 (see after (4.52)) we get that with $\lambda_T = \lambda e^{-\frac{\beta}{4}T}$ we have

$$\begin{aligned}
&P(\alpha_\lambda(\infty) - \alpha_\lambda(T) \leq \lambda\varepsilon/4 | \alpha_\lambda(T) = x)\\
&\qquad\qquad \geq P(\alpha_{\lambda_T}(\infty) \leq \lambda\varepsilon/4 - 2\pi) \geq P(\alpha_{\lambda_T}(\infty) \leq \lambda\varepsilon/8),
\end{aligned}$$

where the second inequality follows if λ is big enough compared to ε. Now we can use part (iii) of Proposition 4.4 with $f(t) = \lambda_T \mathfrak{f}(t)$, $k = 1$ and $a = \lambda\varepsilon/8$ to get

$$P(\alpha_{\lambda_T}(\infty) \le \lambda\varepsilon/8) = 1 - P(\alpha_{\lambda_T}(\infty) > \lambda\varepsilon/8) \ge 1 - 2\frac{8\lambda_T}{2\pi\lambda\varepsilon} \ge 1 - \frac{2}{\pi},$$

where the last step follows from the second assumption of (4.55).

Using this with (4.56) we get that

$$\begin{aligned}\liminf_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \le \varepsilon, t \ge 0)& \\ \ge \liminf_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \le \varepsilon/2, t \in [0,T]),&\end{aligned}$$

and it is enough to estimate the right hand side. We do this by introducing the process $\xi_N(t)$ on $[0,T]$ which is a solution of the SDE (4.7) with initial condition 0 and the piecewise constant drift function $\lambda\mathfrak{f}_N$ where $\mathfrak{f}_N$ is defined as in (4.44). From Proposition 4.4 we have that $\alpha_\lambda(t) \le \xi_N(t)$ and $\widehat{\xi}_N(t) = \xi_N(t) - \alpha_\lambda(t)$ satisfies SDE (4.8) with initial condition 0 and drift $\lambda(\mathfrak{f}_N(t) - \mathfrak{f}(t))$. We have

$$\begin{aligned}P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \le \varepsilon/2, t \in [0,T]) \ge & P(|\lambda^{-1}\xi_N(t) - g(t)| \le \varepsilon/4, t \in [0,T]) \\ & - P\left(\sup_{t\in[0,T]} |\xi_N(t) - \alpha_\lambda(t)| \ge \lambda\varepsilon/4\right).\end{aligned} \tag{4.57}$$

The second term on the right may be bounded in the same manner as (4.48). this gives us

$$\limsup_{\lambda\to\infty} \frac{1}{\lambda^2} \log P\left(\sup_{t\in[0,T]} |\xi_N(t) - \alpha_\lambda(t)| \ge \lambda\varepsilon/4\right) \le -\frac{\beta^2 T^3}{4^2 N^2} \mathcal{I}\left(\frac{\varepsilon N}{\beta T^2}\right).$$

Note, that as $N \to \infty$ the right hand side converges to $-\infty$.

The only thing left is to estimate the first term on the right of (4.57). Introduce the notation $t_k = \frac{Tk}{N}$. We start with the bound

$$\log P(|\lambda^{-1}\xi_N(t) - g(t)| \le \varepsilon/4, t \in [0,T])$$
$$\ge P(|\lambda^{-1}(\xi_N(s+t_k) - \xi_N(t_k)) - (g(s+t_k) - g(t_k))| \le \varepsilon/(4N), s \in [0,T/N]).$$

For any fixed k the process $\xi_N(s+t_k), s \in [0,T/N]$ satisfies the SDE (4.8) with initial condition $\xi_N(t_k)$ and a constant drift $\lambda\mathfrak{f}_N(t_k)$. Using the coupling in the proof of Proposition 4.3 we can construct independent processes $\widehat{\alpha}_k(t), t \in [0,T/N]$ so that

$$\widehat{\alpha}_k(s) - 2\pi \le \xi_N(s+t_k) - \xi_N(t_k) \le \widehat{\alpha}_k(t) + 2\pi, \quad s \in [0,T/N]$$

and $\widehat{\alpha}_k(t), t \in [0,T/N]$ has the same distribution as $\widetilde{\alpha}_{\lambda\mathfrak{f}_N(t_k)}(t), t \in [0,T/N]$. From this it immediately follows that

$$\liminf_{\lambda\to\infty} \tfrac{1}{\lambda^2} \log P(|\lambda^{-1}\xi_N(t) - g(t)| \le \varepsilon/4, t \in [0,T])$$
$$\ge \sum_{k=0}^{N-1} \liminf_{\lambda\to\infty} \tfrac{1}{\lambda^2} \log P(|\lambda^{-1}\widetilde{\alpha}_{\lambda\mathfrak{f}_N(t_k)}(s) - (g(s+t_k) - g(t_k))| < \varepsilon/(8N), s \in [0,T/N]).$$

From our path level large deviation lower bound on $\widetilde{\alpha}$ we get

$$\liminf_{\lambda\to\infty} \tfrac{1}{\lambda^2} \log P(|\lambda^{-1}\widetilde{\alpha}_{\lambda\mathfrak{f}_N(t_k)}(s) - (g(s+t_k) - g(t_k))| \le \varepsilon/(4N), s \in [0,T/N])$$
$$\ge - \inf_{\substack{|\tilde{g}(s)-g(s)|<\varepsilon/(8N)\\ s\in[t_k,t_{k+1}]}} \mathfrak{f}_N(t_k)^2 \int_0^{T/N} \mathcal{I}(\mathfrak{f}_N(t_k)^{-1}\tilde{g}'(t_k+s))ds.$$

This yields the estimate

$$\liminf_{\lambda\to\infty} \tfrac{1}{\lambda^2} \log P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \le \varepsilon/4, t \in [0,T])$$
$$\ge - \inf_{\substack{|\tilde{g}(s)-g(s)|<\varepsilon/(8N),\\ s\in[0,T]}} \int_0^T \mathfrak{f}_N(s)^2 \mathcal{I}(\mathfrak{f}_N^{-1}\tilde{g}'(s))ds.$$

Letting $N \to \infty$ the lower bound converges to $-\int_0^T \mathfrak{f}(s)^2\mathcal{I}(\mathfrak{f}^{-1}(s)g'(s))ds$ which (together with our previous estimates) shows that

$$\liminf_{\lambda\to\infty} \frac{1}{\lambda^2} \log P(|\lambda^{-1}\alpha_\lambda(t) - g(t)| \le \varepsilon, t \ge 0) \ge -\int_0^T \mathfrak{f}(s)^2\mathcal{I}(\mathfrak{f}^{-1}(s)g'(s))ds.$$

Letting $\varepsilon \to 0$ we also have $T = T_\varepsilon \to \infty$ which yields the bound (4.54) and concludes the proof of the lower bound in the large deviation principle. $\square$

4.5 $\mathcal{J}_{\mathbf{Sch},T}$ and $\mathcal{J}_{\mathbf{Sine}_\beta}$ are good rate functions

In this section we will show that $\mathcal{J}_{\mathrm{Sch},T}$ and $\mathcal{J}_{\mathrm{Sine}_\beta}$ are good rate functions. Our main tools are the bound (4.29) and the estimate

$$\mathcal{I}(x) \geq c_1(x-1)^2, \quad \text{if } x > 0 \tag{4.58}$$

both of which will be proved in Proposition (B.2) of the Appendix.

Proposition 4.13. *The functions $\mathcal{J}_{\mathrm{Sine}_\beta}(\cdot)$ and $\mathcal{J}_{\mathrm{Sch},T}(\cdot)$ are both good rate functions on the spaces $C[0,\infty)$ and $C[0,T]$ respectively. Moreover, if $g \in C[0,\infty)$ and $\mathcal{J}_{\mathrm{Sine}_\beta}(g) < \infty$ then $\lim_{t\to\infty} g(t)$ is finite.*

Proof. Fix $T > 0$ and $r \geq 0$. In order to prove that $K_r = \{g : \mathcal{J}_{\mathrm{Sch},T}(\cdot) \leq r\}$ is compact we first show the equicontinuity of this set. Suppose that $g \in K_r$. Then $g(0) = 0$ and $g'(x) \geq 0$ exists a.e. in $[0,T]$. We have for $0 \leq x \leq y \leq T$

$$\begin{aligned} |g(x) - g(y) - (x-y)| &= \left| \int_x^y (g'(s) - 1) ds \right| \\ &\leq (y-x)^{1/2} \sqrt{\int_x^y (g'(s)-1)^2 ds} \\ &\leq c(y-x)^{1/2} \sqrt{\int_x^y \mathcal{I}(g'(s)) ds} \leq c(y-x)^{1/2} r^{1/2} \end{aligned}$$

where we used (4.58) in the second step. This shows that K_r is equicontinuous. Using Tonelli's semicontinuity theorem (e.g. Theorem 3.5, [9]) the compactness of K_r now follows.

The proof for $\mathcal{J}_{\text{Sine}_\beta}(\cdot)$ is bit more involved. Fix $\beta > 0$. It is convenient to transform the interval $[0,\infty)$ into $[0,1)$ using the function $y = 1 - e^{-\beta t/4}$. Then for a $g \in C[0,\infty)$ with $\mathcal{J}_{\text{Sine}_\beta}(g) < \infty$ we have

$$\mathcal{J}_{\text{Sine}_\beta}(g) = \int_0^\infty \mathfrak{f}^2(t)\mathcal{I}(g'(t)\mathfrak{f}^{-1}(t))dt = \frac{\beta}{4}\int_0^1 (1-y)\mathcal{I}(\tilde{g}'(y))dy$$

where $\tilde{g}(y) = g(-\frac{4}{\beta}\log(1-y))$, $\tilde{g} \in C[0,1)$. Consider the functional $\tilde{\mathcal{J}}_{\text{Sine}}(\cdot)$ on $C[0,1)$ defined as

$$\tilde{\mathcal{J}}_{\text{Sine}}(g) = \tfrac{\beta}{4}\int_0^1 (1-t)\mathcal{I}(g'(t))dt \tag{4.59}$$

if $g'(t)$ exists and non-negative for a.e. $0 \le t < 1$, and as ∞ otherwise. Clearly, if we show that $\tilde{\mathcal{J}}_{\text{Sine}}(\cdot)$ is a good rate function on $C[0,1)$ then the same will hold for $\mathcal{J}_{\text{Sine}_\beta}$. We first show that if $g \in C[0,1)$ and $\tilde{\mathcal{J}}_{\text{Sine}}(g) < \infty$ then $\lim\limits_{y\to 1^-} g(y)$ is finite, i.e. we can consider $\tilde{\mathcal{J}}_{\text{Sine}}(\cdot)$ on $C[0,1]$. We have

$$\lim_{y\to 1^-} g(y) = \int_0^1 g'(y)dy \le 2 + \int_0^1 g'(y)\mathbf{1}(g'(y) \ge 2)dy.$$

We will prove that

$$\text{if } h(y) \ge 0, \quad \text{and} \quad \int_0^1 (1-y)h(y)^2\log^2(h(y)+e)dy < \infty, \quad \text{then} \quad \int_0^1 h(y)dy < \infty. \tag{4.60}$$

Using this with $h(y) = g'(y)\mathbf{1}(g'(y) \ge 2)$ together with the bound in (4.29) we get the boundedness of $\int_0^1 g'(y)dy$ and the existence on $\lim\limits_{y\to 1^-} g(y)$.

Let $\Phi(x) = x^2\log^2(|x|+e)$, this is a strictly convex, even function with $\lim\limits_{x\to 0}\frac{\Phi(x)}{x} = 0$, $\lim\limits_{x\to\infty}\frac{\Phi(x)}{x} = \infty$ (i.e. Φ is a 'nice Young function'). Introduce the complementary function

$$\Psi(x) = \Phi^*(x) = \sup_{y\ge 0}\{y|x| - \Phi(y)\} = \int_0^{|x|} (\Phi')^{(-1)}(y)dy$$

where $(\Phi')^{(-1)}$ is the inverse of the strictly increasing function Φ' on $[0,\infty)$. Assume that

$$A = \int_0^1 (1-y)\Phi(h(y))dy < \infty \tag{4.61}$$

and let μ the measure on $[0,1]$ with $d\mu = \frac{1}{A}(1-x)dx$. Consider the Orlicz spaces

$$\begin{aligned} L^\Phi_\mu &= \{f : \text{there is an } a>0 \text{ with } \int_{[0,1]} \Phi(af)d\mu < \infty\}, \\ L^\Psi_\mu &= \{f : \text{there is an } a>0 \text{ with } \int_{[0,1]} \Psi(af)d\mu < \infty\} \end{aligned}$$

with the Luxemburg-norms defined as

$$\|f\|_\Phi = \inf\{b>0 : \int_{[0,1]} \Phi(b^{-1}f)d\mu \le 1\}, \quad \|f\|_\Psi = \inf\{b>0 : \int_{[0,1]} \Psi(b^{-1}f)d\mu \le 1\}. \tag{4.62}$$

(See e.g. [50] for more on Orlicz spaces.) Note, that by our assumption (4.61) we have $\|h\|_\Phi \le 1$. By the generalized Hölder inequality for Orlicz spaces (c.f. Theorem 3 in Chapter III of [50]), for any $f \in L^\Psi_\mu$ one has

$$\|fh\|_1 \le 2\|f\|_\Psi \|h\|_\Phi \le 2\|f\|_\Psi \tag{4.63}$$

where $\|\cdot\|_1$ is the L^1 norm on $[0,1]$ with reference measure μ. Choose $f(x) = \frac{1}{1-x}$. If we show that $\|f\|_\Psi < \infty$ then this would imply

$$\infty > 2\|f\|_\Psi \ge \|fh\|_1 = \tfrac{1}{A}\int_0^1 \frac{1}{1-x} h(x)(1-x)dx = \tfrac{1}{A}\int_0^1 h(x)dx,$$

and the statement (4.60) would follow. It is not hard to check, that there is a $c>0$ so that

$$\Psi(x) \le c\frac{x^2}{\log^2(x+e)}, \qquad \text{for } x \ge 0. \tag{4.64}$$

Since the integral $\int_0^1 (1-x)\frac{(1-x)^{-2}}{\log^2((1-x)^{-1}+e)}dx$ is finite, this implies that $\|\frac{1}{1-x}\|_\Psi$ is finite and thus $\int_0^1 h(y)dy < \infty$. This completes the proof that if $\tilde{\mathcal{J}}_{\text{Sine}}(g) < \infty$ then $\lim_{y\to 1^-} g(y)$ is finite, and also shows the last statement of the proposition.

Next we will prove the equicontinuity of the set $K_r = \{f : \tilde{\mathcal{J}}_{\text{Sine}}(f) \le r\}$, we will show that if $g \in K_r$ then for $\varepsilon < \varepsilon_0$ we have

$$|g(a+\varepsilon) - g(a)| \le C(\log \varepsilon^{-1})^{-1/3} \qquad \text{for any } a \in [0, 1-\varepsilon]. \tag{4.65}$$

Here ε_0, C only depend on r.

We first assume $a \le 1 - \sqrt{\varepsilon}$. Then

$$\begin{aligned}
|g(a+\varepsilon) - g(a) - \varepsilon| &= |\int_a^{a+\varepsilon} (g'(y) - 1)dy| \\
&\le \left(\int_a^{a+\varepsilon} \frac{1}{1-y}dy\right)^{1/2} \left(\int_a^{a+\varepsilon} (1-y)(g'(y)-1)^2 dy\right)^{1/2} \\
&\le Cr^{1/2}\left(\log\left(1 + \frac{\varepsilon}{1-a-\varepsilon}\right)\right)^{1/2} \le Cr^{1/2}\varepsilon^{1/4}
\end{aligned}$$

Where we used $1 - a > \sqrt{\varepsilon}$, the bound (4.58) and the fact that ε can be chosen to be small enough.

Next we assume that $a > 1 - \sqrt{\varepsilon}$. Because of the monotonicity of g it is enough to bound $|g(1) - g(1-\sqrt{\varepsilon})|$. Setting $f(x) = \frac{1}{1-x}$ and $h(x) = g'(x)\mathbf{1}(g'(x) \ge 2)$ we have

$$g(1) - g(1-\sqrt{\varepsilon}) \le 2\sqrt{\varepsilon} + \int_{1-\sqrt{\varepsilon}}^1 h(x)dx. \tag{4.66}$$

Since $\tilde{\mathcal{J}}_{\text{Sine}}(g) \le r$, we can assume that (4.61) holds with some finite $A > 0$. We will now follow the previous argument using Orlicz spaces. We use the same definitions for Ψ, Φ, μ but for the norms $\|\cdot\|_\Psi, \|\cdot\|_\Phi$ defined in (4.62) we use the interval $[1-\sqrt{\varepsilon}, 1]$ instead of $[0, 1]$.

Using inequality (4.66) and (4.63) we get the bound

$$g(1) - g(1-\sqrt{\varepsilon}) \le 2\sqrt{\varepsilon} + A\int_{1-\sqrt{\varepsilon}}^{1} f(x)h(x)d\mu(x) \le 2\sqrt{\varepsilon} + A\|f\|_{\Psi}.$$

To estimate $\|f\|_{\Psi}$ we will prove that with $b = (\log \varepsilon^{-1})^{-1/3}$ there is a constant ε_0 depending on A so that

$$\int_{1-\sqrt{\varepsilon}}^{1} \Psi(b^{-1}f(x))d\mu(x) = A^{-1}\int_{1-\sqrt{\varepsilon}}^{1}(1-x)\Psi(b^{-1}(1-x)^{-1})dx < 1, \qquad \text{for } \varepsilon < \varepsilon_0.$$

This will imply that for such ε we have $\|f\|_{\Psi} \le b$. Using (4.64) we get

$$\begin{aligned} A^{-1}\int_{1-\sqrt{\varepsilon}}^{1}(1-x)\Psi(b^{-1}(1-x)^{-1})dx &\le cb^{-2}A^{-1}\int_0^{\sqrt{\varepsilon}} x^{-1}\frac{1}{\log^2(2x^{-1}b^{-1})}dx \\ &\le \frac{c(\log\varepsilon^{-1})^{2/3}}{A\log\left(2(\log\varepsilon^{-1})^{-1/3}\varepsilon^{-1/2}\right)}. \end{aligned}$$

Since the right hand side converges to 0 as $\varepsilon \to 0$ we get that $\|f\|_{\Psi} \le b$ for small enough ε which in turn leads to the upper bound (4.65). This completes the proof of the equicontinuity of the set K_r and the compactness follows again by Tonelli's theorem. □

4.6 From the path to the endpoint

In this section we will complete the proofs of Theorems 2.9 and 4.1.

Proof of Theorem 4.1. Consider the continuous map $F : C[0,T] \to \mathbb{R}$ given by $F(g) = g(T)/(2\pi)$. By the contraction principle (see e.g. [14]) the random variables $\frac{1}{\lambda}\frac{\alpha_\lambda(T)}{2\pi}$ satisfy a large deviation principle with scale function λ^2 and good rate function J defined as

$$J(\rho) = \min\left\{\int_0^T \mathcal{I}(g'(t))dt : \, g'(t) \ge 0, \, g(T) = 2\pi\rho\right\}. \tag{4.67}$$

We will now solve this variational problem. If g provides the minimum then we can assume that g' is monotone decreasing. To see this define $\tilde{g}$ with $\tilde{g}(0)=0$ and $\tilde{g}'(t)=\sup\{x : m(g'(s)\geq x)\geq t\}$ where m indicates Lesbegue measure. Then $\tilde{g}(T)=g(T)$, $\mathcal{J}_{\mathrm{Sch},T}(\tilde{g})=\mathcal{J}_{\mathrm{Sch},T}(g)$, and $\tilde{g}'(t)$ is decreasing. If $g'>0$ on $[0,a]$ and $g'=0$ on $(a,1]$ then by the classical variational method we get that $\mathcal{I}'(g'(x))$ is constant on $[0,a]$. This means that $g'(x)=\frac{2\pi\rho}{a}$ on $[0,a]$ and $g'(x)=0$ on $(a,T]$ and our variational problem is reduced to finding the minimum of

$$f(a)=a\mathcal{I}\left(\tfrac{2\pi\rho}{a}\right)+(T-a)\mathcal{I}(0), \quad \text{on} \quad 0\leq a\leq T.$$

But we have

$$f'(a)=\mathcal{I}\left(\tfrac{2\pi\rho}{a}\right)-\mathcal{I}(0)-\mathcal{I}'\left(\tfrac{2\pi\rho}{a}\right)\tfrac{2\pi\rho}{a}<0,$$

since $\mathcal{I}$ is strictly convex, which means that the minimum is at $a=T$. Thus

$$J(\rho)=\min\left\{\int_0^T \mathcal{I}(g'(t))dt,\ g'(t)\geq 0,\ g(T)=2\pi\rho\right\}=T\mathcal{I}(2\pi\rho/T) \tag{4.68}$$

is the large deviation rate function for $\frac{1}{\lambda}\frac{\alpha_\lambda(T)}{2\pi}$.

Now recall that the counting function of Sch_τ is given by

$$\widetilde{N}_\tau(\lambda)=\#\{\nu : 0\leq\nu\leq\lambda, \phi_{\lambda/\tau}(\tau)\in 2\pi\mathbb{Z}\}$$

where ϕ_λ is the solution of (4.2). Note, that $\phi_\lambda(t)-\phi_0(t)$ has the same distribution as $\xi_{f,0}(t)$ with constant $f=\lambda$, which in turn has the same distribution as $\widetilde{\alpha}_\lambda(t)$. Using the coupling methods of Proposition 4.4 we can show that $\phi_\lambda(t)$ is increasing in λ for any fixed t (see [43] for a detailed proof of this fact). From this it follows that

$$\left|\tilde{N}_\tau(\lambda)-\tfrac{1}{2\pi}\left(\phi_{\lambda/\tau}(\tau)-\phi_0(\tau)\right)\right|\leq 1.$$

This means that in order to get a large deviation principle for $\frac{1}{\lambda}\tilde{N}_\tau(\lambda)$ it is enough to prove one for $\frac{1}{\lambda}\frac{\phi_{\lambda/\tau}(\tau)-\phi_0(\tau)}{2\pi}$. But this has the same distribution as $\frac{1}{\lambda}\frac{\widetilde{\alpha}_{\lambda/\tau}(\tau)}{2\pi}$, and a simple rescaling of (4.68) completes the proof of the theorem. □

Proof of Theorem 2.9. Theorem 4.5 shows that $\frac{1}{\lambda}\alpha_{\lambda(\cdot)}$ satisfies a path level large deviation principle. By applying the time change $y = 1-e^{-\frac{\beta}{4}t}$, we get that $t \to \frac{1}{\lambda}\alpha_\lambda(1-e^{-\frac{\beta}{4}t})$ satisfies a path level LDP on $C[0,1)$ with the modified rate function $\tilde{\mathcal{J}}_{\text{Sine}}$ given in (4.59). In Proposition 4.13 we showed that if $\tilde{\mathcal{J}}_{\text{Sine}}(g) < \infty$ then the limit as $t \to 1^-$ exists and so the LDP actually holds on $C[0,1]$. Using the contraction principle with the functional $F(g) = \frac{1}{2\pi}g(1)$, we get that $\frac{1}{\lambda}\frac{\alpha_\lambda(\infty)}{2\pi}$ satisfies a large deviation principle with speed function λ^2 and a good rate function

$$\begin{aligned} J^\beta(\rho) &= \min\left\{\tilde{\mathcal{J}}_{\text{Sine}}(g) :\, g(1) = 2\pi\rho\right\} \\ &= \min\left\{\tfrac{\beta}{4}\int_0^1 (1-t)\mathcal{I}(g'(t))dt :\, g(0) = 0, g'(t) \geq 0, g(1) = 2\pi\rho\right\}. \end{aligned}$$

The counting function $N_\beta(\lambda)$ of Sine_β is given by $\frac{\alpha_\lambda(\infty)}{2\pi}$, so Theorem 2.9 will follow if we can show that the solution of this variational problem is given by $\beta I_{\text{Sine}}(\rho)$ as defined in the theorem.

The function $\tilde{\mathcal{J}}_{\text{Sine}}$ is a good rate function, so for any $\rho \geq 0$ the minimum is achieved at some $g_\rho \in C[0,1]$. Clearly, when $\rho = \frac{1}{2\pi}$ then the minimum is zero, as the $g(t) = t$ function shows. (We will not denote the dependence of ρ in $g = g_\rho$ from this point.)

We may assume that for the minimizer the derivative g' will not take values from both $(1,\infty)$ and $[0,1)$ because otherwise we could construct a function $\hat{g}$ with the same boundary condition $\hat{g}(1) = 2\pi\rho$, but with $\tilde{\mathcal{J}}_{\text{Sine}}(g) > \tilde{\mathcal{J}}_{\text{Sine}}(\hat{g})$. The construction is as follows. Assume $\rho < 1/(2\pi)$ and that $A = \{t : g'(t) > 1\}$ has positive measure. Since $\int_0^1 (g'(t)-1)dt = 2\pi\rho - 1 < 0$ and $\int_A (g'(t)-1)dt > 0$, by the intermediate value theorem we can find $B \subset [0,1] \setminus A$ so that $\int_{A\cup B}(g'(t)-1)dt = 0$. Define $\hat{g}$ with $\hat{g}(0)=0$, $\hat{g}'(t) = g'(t)$ if $t \notin A \cup B$ and $\hat{g}'(t) = 1$ otherwise. Then $g(1) = \int_0^1 g'(t)dt = \int_0^1 \hat{g}'(t)dt = \hat{g}(1)$, but clearly $\tilde{\mathcal{J}}_{\text{Sine}}(g) > \tilde{\mathcal{J}}_{\text{Sine}}(\hat{g})$. A similar construction works for $\rho > 1/(2\pi)$. Thus we may assume that $g'(t) \le 1$ for all t if $\rho < 1/(2\pi)$, and $g'(t) \ge 1$ for all t if $\rho > 1/(2\pi)$.

First assume that $\rho > \frac{1}{2\pi}$. Then $g'(t) \ge 1$ for all t and we can use the classical variational method (see e.g. [9]) to conclude that $(1-t)\mathcal{I}'(g'(t))$ is constant in t. Thus the optimizer is given by a function g_ρ which satisfies

$$g_\rho(0) = 0, \qquad \mathcal{I}'(g'(t)) = \tfrac{c_\rho}{1-t}, \qquad \int_0^1 (\mathcal{I}')^{(-1)}\left(\tfrac{c_\rho}{1-t}\right) dt = 2\pi\rho, \tag{4.69}$$

for some constant c_ρ and the solution of the variational problem is

$$J^\beta(\rho) = \tfrac{\beta}{4}\int_0^1 (1-t)\mathcal{I}\left((\mathcal{I}')^{(-1)}\left(\tfrac{c_\rho}{1-t}\right)\right) dt. \tag{4.70}$$

In Proposition 4.14 below we will show that this is equal to $\beta I_{Sine}(\rho)$ as defined in Theorem 2.9.

Now assume that $\rho < \frac{1}{2\pi}$, here we can assume that the minimizer satisfies $g'(t) \le 1$. As in the case of Sch_τ we may assume g' is decreasing, this can be shown using the same construction as found in the paragraph directly following equation (4.67). Suppose that g' is zero for $t \in [a,1]$ and $g'(t) > 0$ in $[0,a]$. Then on $[0,a]$ the classical variational method shows that $(1-t)\mathcal{I}'(g'(t))$ must be constant. Thus the optimizer must be of the following form:

$$g'(t) = \begin{cases} (\mathcal{I}')^{(-1)}\left(\tfrac{c_{\rho,a}}{1-t}\right), & 0 \le t \le a \\ 0, & a < t \le 1, \end{cases} \tag{4.71}$$

for some constant c_ρ which satisfies

$$2\pi\rho = \int_0^a (\mathcal{I}')^{(-1)}\left(\tfrac{c_{\rho,a}}{1-t}\right) dt. \tag{4.72}$$

By Propositions B.1 and B.2 of the Appendix the function $\mathcal{I}'(x)$ is strictly increasing on $(0,\infty)$ with a limit of $-\frac{1}{2\pi}$ at $x=0$. Thus $c_{\rho,a}$ in (4.71) cannot be smaller than $-\frac{1-a}{2\pi}$. Our next claim is that the optimizer has a continuous derivative at $t=a$, which will identify $c_{\rho,a}$ as $c=-\frac{1-a}{2\pi}$. Assume the opposite, i.e. that $c_{\rho,a}>-\frac{1-a}{2\pi}$ and $g'(a)>0$. Let $\eta_\delta(x)=\mathbf{1}_{(a,a+\delta)}-\mathbf{1}_{(a-\delta,a)}$. If δ,ε are small enough then $g'-\varepsilon\eta_\delta\ge 0$ in $[0,1]$ and $\tilde g(t)=\int_0^t(g'(s)-\varepsilon\eta_\delta(s))ds$ satisfies the same boundary conditions as g. Since g is a minimizer, the derivative of $h(\varepsilon)=\tilde{\mathcal{J}}_{\text{Sine}}(g+\varepsilon\eta_\delta)$ at $\varepsilon=0$ cannot be negative. We can compute the derivative as

$$h'(0)=\frac{\beta}{4}\int_0^1(1-t)\mathcal{I}'(g(t))\eta_\delta(t)dt=-\int_{a-\delta}^a(1-t)\tfrac{c_{\rho,a}}{1-t}dt+\int_a^{a+\delta}(1-t)\left(-\tfrac{1}{2\pi}\right)dt.$$

This is equal to $\delta(-c_{\rho,a}-\frac{1-a}{2\pi})+\frac{\delta^2}{4\pi}$ which is negative if δ is small enough (by our assumption that $c_{\rho,a}>-\frac{1-a}{2\pi}$). The contradiction shows that we must have $c=-\frac{1-a}{2\pi}$.

Thus the optimizer is given by

$$g'(t)=\begin{cases}(\mathcal{I}')^{(-1)}\left(\frac{a-1}{2\pi(1-t)}\right), & 0\le t\le a\\ 0, & a<t\le 1.\end{cases}\tag{4.73}$$

for some $0\le a\le 1$ with

$$2\pi\rho=\int_0^a(\mathcal{I}')^{(-1)}\left(\tfrac{a-1}{2\pi(1-t)}\right)dt.\tag{4.74}$$

and the solution of the variational problem in the $2\pi\rho<1$ case is given by

$$J^\beta(\rho)=\frac{\beta}{4}\int_0^a(1-t)\mathcal{I}\left((\mathcal{I}')^{(-1)}\left(\tfrac{a-1}{2\pi(1-t)}\right)\right)dt+\tfrac{\beta}{64}(1-a)^2.\tag{4.75}$$

In Proposition 4.14 below we will show that this is equal to $\beta I_{\text{Sine}}(\rho)$. □

Proposition 4.14. *The rate function for the* Sine_β *process is given by*

$$\beta I_{Sine}(\rho)=\frac{\beta}{8}\left[\frac{\nu}{8}+\rho\mathcal{H}(\nu)\right]$$

where $\nu = \gamma^{-1}(\rho)$, and γ is the strictly increasing function given in (2.13).

Proof. We have to show that $J^\beta(\rho)$ defined by (4.69) and (4.70) for $\rho > 1/(2\pi)$ and by (4.74) and (4.75) for $\rho < 1/(2\pi)$ is equal to βI_{Sine} given above.

We begin with the case where $\rho > \frac{1}{2\pi}$. In this case the minimizer $g = g_\rho$ is given by (4.69). One easily checks that

$$\frac{d}{dt}\left(\tfrac{\beta}{8}\left(-(1-t)^2\mathcal{I}(g'(t)) + c_\rho(1-t)g'(t) + c_\rho g(t)\right)\right) = \tfrac{\beta}{4}(1-t)\mathcal{I}(g'(t)). \tag{4.76}$$

From this we get

$$J^\beta(\rho) = \frac{\beta}{4}\int_0^1 (1-t)\mathcal{I}\left(g'(t)\right) dt = \tfrac{\beta}{8}\left[\mathcal{I}(g'(0)) - c_\rho g'(0) + 2\pi\rho c_\rho\right]$$

where we used $g(0) = 0$, $g(1) = 2\pi\rho$, and the limits

$$\lim_{t\to 1^-}(1-t)^2\mathcal{I}(g'(t)) = \lim_{x\to\infty}\frac{c_\rho^2\mathcal{I}(x)}{\mathcal{I}'(x)^2} = 0, \qquad \lim_{t\to 1^-}(1-t)g'(t) = \lim_{x\to\infty}\frac{x}{c_\rho\mathcal{I}'(x)} = 0$$

which follow from the asymptotics (B.6) and (B.7) to be proven in Proposition B.2.

Now for the case where $\rho < \frac{1}{2\pi}$ we have that g_ρ is given by (4.73). Using the notation $c = c_\rho = \frac{a-1}{2\pi}$, the identity (4.76) gives

$$J^\beta(\rho) = \frac{\beta}{4}\int_0^{2\pi c_\rho+1}(1-t)\mathcal{I}\left(g'(t)\right) dt + \tfrac{\beta}{8}(2\pi c_\rho)^2\mathcal{I}(0) = \tfrac{\beta}{8}\left[\mathcal{I}(g'(0)) - c_\rho g'(0) + 2\pi\rho c_\rho\right],$$

where we used $g(0) = 0$, $g(a) = 2\pi\rho$, and $g'(a) = 0$. Note, that $c_\rho > 0$ if $\rho > 1/(2\pi)$ and $-\frac{1}{2\pi} \le c_\rho < 0$ if $\rho < 1/(2\pi)$. Introducing $\nu = K^{(-1)}\left(\frac{\pi}{2(\mathcal{I}')^{(-1)}(c_\rho)}\right)$, we get for both $\rho < 1/(2\pi)$ and $\rho > 1/(2\pi)$ that

$$J^\beta(\rho) = \frac{\beta}{8}\left(\frac{\nu}{8} + \rho\mathcal{H}(\nu)\right)$$

which agrees with (2.15), we just have to show that $\nu = \gamma^{(-1)}(\rho)$. Note, that $\nu = \nu(\rho) < 0$ if $2\pi\rho > 1$ and $0 < \nu < 1$ if $2\pi\rho < 1$.

Recall from (4.69) and (4.74) that

$$\rho = \tfrac{1}{2\pi}\int_0^1 (\mathcal{I}')^{(-1)}\left(\tfrac{c_\rho}{1-t}\right)dt, \quad \text{if } \rho > \frac{1}{2\pi}, \qquad \text{and} \qquad \rho = \tfrac{1}{2\pi}\int_0^{2\pi c+1} (\mathcal{I}')^{(-1)}\left(\tfrac{c_\rho}{1-t}\right)dt, \quad \text{if } \rho < \frac{1}{2\pi}.$$

Applying the change of variables to both integrals with a new variable x satisfying $\frac{\pi}{2K(x)} = (\mathcal{I}')^{(-1)}(\frac{c_\rho}{1-t})$, we get that ρ depends on $\nu = K^{(-1)}\left(\frac{\pi}{2(\mathcal{I}')^{(-1)}(c_\rho)}\right)$ exactly via (2.13) which finishes the proof. Note, that the finiteness of the integrals in (2.13) follow from the asymptotics of $K(x)$ and $E(x)$ near 1 and $-\infty$ (see the proof of Proposition B.2). $\square$

Chapter 5

Large Deviations and a Central Limit Theorem for the $\text{Bess}_{a,\beta}$ Process

The proof of the large deviation and the CLT for $\text{Bess}_{a,\beta}$ rests on the fact that the distribution of the number of points in a large interval can be described with a stochastic differential equation similar to that that defines the counting function $N_\beta(\lambda)$ of the Sine_β process. In both cases, this new characterization allows us to use the ideas used for the Sine_β process.

To begin we need to characterize the $\text{Bess}_{a,\beta}$ process using diffusions rather then an operator. We consider the 'Riccati diffusion' for $\mathfrak{G}_{\beta,a}$, given by the stochastic differential equation

$$dp(t) = \frac{2}{\sqrt{\beta}} p(t) dB(t) + \left(\left(a + \tfrac{2}{\beta} \right) p(t) - p^2(t) - \lambda e^{-t} \right) dt, \tag{5.1}$$

with initial condition $p(0) = +\infty$, which it leaves instantaneously. Note that there is a positive probability of explosion to $-\infty$.

Theorem 5.1 ([48]). *Let $\Lambda_0(\beta, a) < \Lambda_1(\beta, a) < \ldots$ be the ordered eigenvalues of $\mathfrak{G}_{\beta,a}$, and let $P_{\infty,t}$ denote the law induced by $p(\cdot : \beta, a, \lambda)$ started at $+\infty$ at time t, and restarted* *at* $+\infty$ *and time* $\mathfrak{m}$ *upon any* $\mathfrak{m} < \infty, p(\mathfrak{m}) = -\infty$. *Then,*

$$P(\Lambda_0(\beta, a) > \lambda) = P_{\infty,0}(p \text{ never hits } 0), \tag{5.2}$$

$$P(\Lambda_k(\beta, a) < \lambda = P_{\infty,0}(p \text{ hits 0 at least } k+1 \text{ times}). \tag{5.3}$$

The remainder of this chapter will be organized as follows: Section 5.1 will give another SDE characterization of the $\text{Bess}_{a,\beta}$ process similar to that in the Sine_β case. Section 5.2 will give a proof of the central limit theorem using the characterization in terms of the new diffusion. Section 5.3 will give and prove an LDP for the SDE from the preceding section. The final section will use this path level large deviation to derive a large deviation principle for the number of points in a large interval of $\text{Bess}_{a,\beta}$.

5.1 Another characterization

In order to use the work we did on the Sine_β process we will rewrite the SDE characterization of the $\text{Bess}_{a,\beta}$ process with different diffusion and give the distribution of the number of points in an interval.

Theorem 5.2. *Let* $M_{a,\beta}(\lambda)$ *be the number of points of* $Bess_{a,\beta}$ *in the interval* $[0, \lambda]$, *and let* $\varphi_{a,\lambda}$ *be the diffusion that satisfies the SDE*

$$d\varphi_{a,\lambda} = \frac{\beta}{2}(a + 1/2)\sin\left(\frac{\varphi_{a,\lambda}}{2}\right)dt + \beta\sqrt{\lambda}e^{-\beta t/8}dt + \frac{\sin\varphi_{a,\lambda}}{2}dt - 2\sin\left(\tfrac{\varphi_{a,\lambda}}{2}\right)dB_t \tag{5.4}$$

with initial condition $\varphi_{a,\lambda}(0) = 2\pi$. *Then*

$$M_{a,\beta}(\lambda) =^d \lim_{t\to\infty}\left\lfloor \frac{1}{4\pi}\varphi_{a,\lambda}(t)\right\rfloor.$$

proof of Theorem 5.2. The characterization of the $\mathrm{Bess}_{a,\beta}$ process given in Theorem 5.1 can be rewritten in the following way: We can apply the change of variables $-X_1(t) := \log(p(\beta t/4)) + \beta t/8 - \log\lambda/2$ for $p > 0$. Then X_1 satisfies the SDE

$$dX_1(t) = \left(\frac{\beta}{4}(-a - \tfrac{1}{2}) + \frac{\beta}{2}\sqrt{\lambda}e^{-\beta t/8}\cosh X(t)\right)dt + dB(t) \tag{5.5}$$

The initial condition $p(0) = +\infty$ give $X(0) = -\infty$, moreover when $p(\beta t/4)$ reaches 0 we get that $X(t) = +\infty$.

Notice that this change of variables is for $p > 0$, we can do a similar change of variables with $p < 0$, where we take $X_2(t) = \log(-p(\beta t/4)) + \beta t/8 - \log\lambda/2$. This gives us

$$dX_2(t) = \left(\frac{\beta}{4}(a + \tfrac{1}{2}) + \frac{\beta}{2}\sqrt{\lambda}e^{-\beta t/8}\cosh X(t)\right)dt + dB(t). \tag{5.6}$$

When $p(\beta t/4) = 0$ this gives $X_2(t) = -\infty$, and when $p(\beta t/4) = -\infty$ this gives $X_2(t) = +\infty$.

To find the $\varphi_{a,\lambda}$ diffusion given in Theorem 5.2 we work back from X_1 and X_2. Let $-\log\tan(-\varphi/4) = X_1$, then $\varphi = -4\arctan e^{-X_1}$ and we get

$$\begin{aligned} d\varphi &= -2 \text{ sech } X_1 dX_1 + \text{ sech } X_1 \tanh X_1 dt \\ &= \frac{\beta}{2}(a + \tfrac{1}{2})\sin\left(\frac{\varphi}{2}\right)dt + \beta\sqrt{\lambda}e^{-\beta t/8}dt + \frac{\sin\varphi}{2}dt - 2\sin\left(\frac{\varphi}{2}\right)dB_t. \end{aligned}$$

The conditions $X_1 = -\infty$ and $X_1 = +\infty$ correspond to $\varphi = -2\pi$ and $\varphi = 0$ respectively.

Now lets move on to X_2. Then take $\varphi = 4\arctan e^{X_2}$. This gives us

$$\begin{aligned} d\varphi &= -2 \text{ sech } X_2 dX_2 + \text{ sech } X_2 \tanh X_2 dt \\ &= \frac{\beta}{2}(a + 1/2)\sin\left(\tfrac{\varphi}{2}\right)dt + \beta\sqrt{\lambda}e^{-\beta t/8}dt + \frac{\sin\varphi}{2}dt + 2\sin\left(\tfrac{\varphi}{2}\right)dB_t. \end{aligned}$$

The conditions $X_2 = -\infty$ and $X_2 = +\infty$ correspond to $\varphi = 0$ and 2π respectively. Then since B_t is equal in distribution to $-B_t$ these describe the same diffusion.

Now notice that this diffusion is invariant under 4π spacial shifts, so for a fixed λ with initial condition $\varphi_{a,\lambda}(0) = 2\pi$ then

$$P(\sup_t \varphi_{a,\lambda}(t) \geq 4\pi k) = P(\Lambda_k < \lambda).$$

$\square$

5.2 The central limit theorem

The proof of the central limit theorem is similar to the proof for the Sine_β process which was done by Kritchevski, Valkó, and Virág in [43].

Proof of Theorem 2.10. First notice that the process $\hat{\varphi}_{a,\lambda}(t) = \varphi_{a,\lambda}(t+T)$ with $T = \frac{8}{\beta}\log(\sqrt{\lambda})$ satisfies the same SDE with $\lambda = 1$. This gives us that

$$\frac{\varphi_{a,\lambda}(\infty) - \varphi_{a,\lambda}(T)}{\sqrt{\log \lambda}} \to 0$$

in probability. From this we have that it is sufficient to consider the weak limit of

$$\frac{\varphi_{a,\lambda}(T) - 8\sqrt{\lambda}}{4\pi\sqrt{\log \lambda}}.$$

Written in its integrated form, the SDE for $\varphi_{a,\lambda}$ gives us that

$$\varphi(T) - 8\sqrt{\lambda} + 8 - 2\pi = \frac{\beta}{2}(a + 1/2)\int_0^T \sin\left(\frac{\varphi}{2}\right)dt + \int_0^T \frac{\sin(\varphi)}{2}dt + 2\int_0^T \sin\left(\tfrac{\varphi}{2}\right)dB_t. \tag{5.7}$$

We will show that when scaled down by $\log \lambda$ the first two terms vanish in the limit, then show that the noise term gives the appropriate contribution. Using Itô's equation we get that

$$\frac{d\left(e^{i\varphi+\beta t/8}\right)}{i\sqrt{\lambda}\log\lambda} = \tfrac{1}{\sqrt{\lambda}\log\lambda}e^{i\varphi+\beta t/8}d\varphi + \tfrac{\beta}{8i\sqrt{\lambda}\log\lambda}e^{i\varphi+\beta t/8}dt - \tfrac{2}{i\sqrt{\lambda}\log\lambda}e^{i\varphi+\beta t/8}\sin^2\left(\tfrac{\varphi}{2}\right)dt$$

$$= \tfrac{\beta}{\sqrt{\log\lambda}}e^{i\varphi}dt + \tfrac{2}{\sqrt{\lambda}\log\lambda}\sin\left(\tfrac{\varphi}{2}\right)e^{i\varphi+\beta t/8}dB_t$$

$$+ \tfrac{1}{\sqrt{\lambda}\log\lambda}\left[\frac{\beta}{2}(a+\tfrac{1}{2})\sin\left(\tfrac{\varphi}{2}\right) + \frac{\sin\varphi}{2} - i\frac{\beta}{8} + i2\sin^2\left(\tfrac{\varphi}{2}\right)\right]e^{i\varphi+\beta t/8}dt \tag{5.8}$$

Integrated, the left hand side gives us $\frac{1}{i\sqrt{\log\lambda}}[e^{i\varphi(T)}-1]$ which is order $O((\log\lambda)^{-1/2})$. The integral of line (5.8) is of order $\frac{1}{\sqrt{\lambda}\log\lambda}\int_0^T e^{i\varphi+\beta t/8}dt = O((\log\lambda)^{-1/2})$. Lastly the integral of the noise term is has an L^2 norm bounded by $C(\log\lambda)^{-1/2}$. This gives us that $(\log\lambda)^{-1/2}\int_0^T e^{i\varphi}dt \to 0$ in probability as $\lambda\to\infty$ which shows that the second integral term in (5.7) will not contribute when scaled by . This shows that we can do a similar calculation using $e^{i\varphi/2+\beta t/8}$ to get that the first integral term on the right hand side of (5.7) also makes no contribution.

We now turn our attention to the last remaining term in (5.7). We can rewrite this as

$$2\int_0^T \sin\left(\tfrac{\varphi}{2}\right)dB_t = \hat{B}\left(4\int_0^T \sin^2\left(\tfrac{\varphi}{2}\right)dt\right)$$

for some standard Brownian motion $\hat{B}_t$. Then we have that

$$\frac{4}{\log\lambda}\int_0^T \sin^2\left(\tfrac{\varphi}{2}\right)dt = \frac{16}{\beta} + \frac{2}{\log\lambda}\int_0^T \cos\left(\tfrac{\varphi}{2}\right)dt. \tag{5.9}$$

This final integral term goes to 0 in probability from the work done earlier that showed $(\log\lambda)^{-1/2}\int_0^T e^{i\varphi}dt \to 0$ in probability as $\lambda\to\infty$. This completes the proof.

$\square$

5.3 An LDP for $\varphi_{a,\lambda}$

As in the case of the Sine$_\beta$ process we begin by proving an LDP for the diffusion that characterizes the counting function of the process.

Theorem 5.3. *Fix $\beta > 0$ and let $\varphi_{a,\lambda}(t)$ be the process defined in (5.4) with $a > -1/2$. Then the sequence of rescaled processes $(\frac{\varphi_{a,\lambda}(t)}{\sqrt{\lambda}}, t \in [0,\infty))$ satisfies a large deviation principle on $C[0,\infty)$ with scale λ and good rate function $\mathcal{J}_{Bess_{a,\beta}}$. The rate function $\mathcal{J}_{Bess_{a,\beta}}$ is defined as*

$$\mathcal{J}_{Bess_{a,\beta}}(g) = \int_0^\infty \mathfrak{h}^2(t)\mathcal{I}\left(g'(t)/\mathfrak{h}(t)\right) dt, \quad \text{with} \quad \mathfrak{h}(t) = \mathfrak{h}_\beta(t) = \beta e^{-\frac{\beta}{8}t}$$

in the case where $g(0) = 2\pi$ and g is absolutely continuous with non-negative derivative g'. In all other cases $\mathcal{J}_{Bess_{a,\beta}}(g)$ is defined as ∞.

Proof. The proof of this theorem will follow exactly the proof of Theorem 4.5 once we have the following two pieces:

Let $\tilde{\varphi}_{a,\lambda}$ be the diffusion that satisfies

$$d\tilde{\varphi}_{a,\lambda} = \frac{\beta}{2}(a+1/2)\sin\left(\tfrac{\tilde{\varphi}_{a,\lambda}}{2}\right)dt + \beta\sqrt{\lambda}dt + \frac{\sin\tilde{\varphi}_{a,\lambda}}{2}dt + 2\sin\left(\tfrac{\tilde{\varphi}_{a,\lambda}}{2}\right)dB_t, \tag{5.10}$$

with initial condition $\tilde{\varphi}_{a,\lambda}(0) = 2\pi$. We will also use the corresponding constant λ versions of X_1 and X_2. Take

$$d\tilde{X}_1 = \left(\frac{\beta}{4}(-a-\tfrac{1}{2}) + \sqrt{\lambda}\cosh\tilde{X}_1(t)\right)dt + dB_t \tag{5.11}$$

$$d\tilde{X}_2 = \left(\frac{\beta}{4}(a+\tfrac{1}{2}) + \sqrt{\lambda}\cosh\tilde{X}_2(t)\right)dt + dB_t \tag{5.12}$$

Proposition 5.4. *Let $\tau_1 = \inf\{t : \tilde{\varphi}_{a,\lambda}(t) = 4\pi$, and $\tau_2 = \inf\{t : \tilde{\varphi}_{a,\lambda} = 6\pi\} - \tau_1$ and fix $A < 1$, then for $\tau = \tau_1$ or τ_2 we have*

$$Ee^{\frac{\lambda A}{8}\tau - \frac{\sqrt{\lambda}\tau}{4}(|A|\wedge\sqrt{|A|})(1+\frac{\beta}{2}(a+\frac{1}{2}))} \le e^{-\sqrt{\lambda}\mathcal{H}(A)}. \tag{5.13}$$

Let $t_A = 4K(A)$ and fix $0 < \varepsilon < |t_A - 2\pi|$. Then we have

$$P(\sqrt{\lambda}\tau \in [t_A - \varepsilon, t_A + \varepsilon]) \geq C(\varepsilon, \lambda, A) e^{-\sqrt{\lambda}(\mathcal{H}(A) + \frac{At_A}{8}) - \sqrt{\lambda}\frac{|A|\varepsilon}{8} - \sqrt{\lambda}\frac{|A|}{2}(t_A+\varepsilon)(1+\frac{\beta}{2}(a+\frac{1}{2}))} \tag{5.14}$$

where $\lim_{\lambda\to\infty} C(\varepsilon, \lambda, A) = 1$ *for fixed* a, ε.

The proof of this proposition follows the same steps as the proof of Proposition 4.7, and will not be shown.

The other piece that we need to complete the proof of the path LDP for $\varphi_{a,\lambda}$ is a tail bound.

Proposition 5.5. *Let $\varphi_{a,\lambda}$ be the diffusion defined in (5.4), $T > 0$ and $\lambda > 5\pi/\epsilon$, then for $a > -1$*

$$\lim_{\lambda\to\infty} \frac{1}{\lambda} \log P(\varphi_{a,\lambda}(\infty) - \varphi_{a,\lambda}(T\sqrt{\lambda}) \geq \epsilon\sqrt{\lambda}) \leq -\frac{\epsilon\beta^2}{2^8\pi} T \max\{a + \tfrac{1}{2}, -a - \tfrac{1}{2}\}. \tag{5.15}$$

Notice that this gives us the necessary upper bound on the tail in the case where $a \neq -1/2$.

These two pieces are sufficient to complete the proof of the path deviation for the $\varphi_{a,\lambda}$ in the same way as the proof of Theorem 4.5.

$\square$

Proof of Proposition 5.5. The idea of the proof will be to consider the process on the intervals of the type $[4k\pi - 2\pi, 4k\pi]$ when $a > -1/2$, and on intervals of the type $[4k\pi, 4k\pi + 2\pi]$ for $-1 < a < -1/2$. The diffusion $\varphi_{a,\lambda}$ will need to cross at least $\lfloor \epsilon\sqrt{\lambda}/(4\pi) \rfloor$ many of these, therefore

$$P(\varphi_{a,\lambda}(\infty) - \varphi_{a,\lambda}(T\sqrt{\lambda}) \geq \epsilon\sqrt{\lambda}) \leq P(\tilde{\varphi}_{a,\lambda\mathfrak{g}^2(T\sqrt{\lambda})}(\infty) \geq 2\pi)^{\lfloor \epsilon\sqrt{\lambda}/(4\pi) \rfloor} \tag{5.16}$$

where $\tilde{\varphi}(0) = 0$ when $a \in (-1, -1/2)$, and $\tilde{\varphi}(0) = 2\pi$ when $a > -1/2$. Now to get an upper bound on this probability we need to look at the $\tilde{X}_2$ diffusion defined in (5.12) for $a > -1/2$, and the $\tilde{X}_1$ diffusion for $a \in (-1, -1/2)$. We will compare the diffusions to a reflecting Brownian Motion with drift. We begin with the $\tilde{X}_2$ diffusion, let $Z_{a,\lambda}$ be a Brownian motion with drift $\frac{\beta}{4}(-a - \frac{1}{2}) + \frac{\beta}{2}\sqrt{\lambda}e^{-\beta T\sqrt{\lambda}/16}$ reflected at the origin, and let $\tau_{Z,x}$ be the hitting time of x for Z. Then

$$
\begin{aligned}
P(\tilde{\varphi}_{a,\lambda\mathfrak{g}^2(T\sqrt{\lambda})}(\infty) \geq 2\pi) &\leq P(\tau_{Z,\beta T\sqrt{\lambda}/16} < \infty) \\
&\leq \exp\left[\left(\frac{\beta}{4}(-a - \tfrac{1}{2}) + \frac{\beta}{2}\sqrt{\lambda}e^{-\beta T\sqrt{\lambda}/16}\right)\frac{\beta}{16}T\sqrt{\lambda}\right]
\end{aligned}
$$

where the last bound is a standard result on reflected Brownian Motion (see e.g. [30] and [39]). A similar bound can be used on the $\tilde{X}_1$ diffusion using a reflecting Brownian Motion with drift $\frac{\beta}{4}(a + \frac{1}{2}) + \frac{\beta}{2}\sqrt{\lambda}e^{-\beta T\sqrt{\lambda}/16}$. These bounds together with (5.16) are sufficient to complete the proof.

□

5.4 From the path to the endpoint

This part of the proof also follows nearly immediately from the work done in the Sine_β case. Consider the $\mathcal{J}_{\text{Bess}_{a,\beta}}$ rate function given in Theorem 5.3, we apply the change of variables $x = (1 - e^{-\frac{8}{\beta}t})$ to get the modified rate function

$$
\tilde{\mathcal{J}}_{\text{Bess}_{a,\beta}}(\tilde{g}) = 8\beta \int_0^1 (1 - x)\mathcal{I}(\tilde{g}'(x)/8)dy
$$

where $\tilde{g}(x) = g(-\frac{8}{\beta}\log(1 - x))$. In other words

$$
\tilde{\mathcal{J}}_{\text{Bess}_{a,\beta}}(g) = 32\tilde{\mathcal{J}}_{\text{Sine}}((g - 2\pi)/8).
$$

Notice that we have a shift of 2π because we assume $\varphi_{a,\lambda}(0) = 2\pi$ whereas $\alpha_\lambda(0) = 0$ Lastly recall that we want to optimize over g with endpoint $4\pi\rho$ where in the Sine_β case we had an endpoint of $2\pi\rho$. This gives us that $I_{\mathrm{Bess}_{a,\beta}}(\rho) = 32 I_{\mathrm{Sine}}(\rho/4)$.

Appendix A

Proof of Theorem 3.1

We will be working primarily with a lower bidiagonal matrix M, and the symmetric tridiagonal matrix MM^T. For convenience we will adopt the following notations. Denote the diagonal entries of our symmetric tridiagonal matrix MM^T by $\underline{a} = \{a_1, ..., a_n\}$ and its off-diagonal entries by $\underline{b} = \{b_1, ..., b_{n-1}\}$. For the entries of the bi-diagonal matrix M use $\underline{x} = \{x_1, ..., x_n\}$ to denote the diagonal entries and $\underline{y} = \{y_1, ..., y_{n-1}\}$ for its sub-diagonal entries. We will also use another tridiagonal matrix which arises in the following lemma (see e.g. [18]):

Lemma A.1. *Let M be an $n \times n$ bidiagonal matrix with $a_1, a_2, \ldots, a_n$ in the diagonal and $b_1, b_2, \ldots, b_{n-1}$ in the off-diagonal. Consider the $2n \times 2n$ symmetric tridiagonal matrix L which has zeros in the main diagonal and $a_1, b_1, a_2, b_2, \ldots, a_n$ above and below the diagonal. If the singular values of M are $\lambda_1, \lambda_2, \ldots, \lambda_n$ then the eigenvalues of L are $\pm\lambda_i, i = 1 \ldots n$.*

Recall the definitions of $M_{\beta,n}, C_k$ and $\tilde{C}_k$ from the statement of Theorem 3.1. It will be convenient write $\{C_i, S_i\}_{i=1}^{n}$ and $\{\tilde{C}_i, \tilde{S}_i\}_{i=1}^{n-1}$ in terms of a set of random angles $\underline{\alpha} = \{\alpha_1, ..., \alpha_n\}$ and $\underline{\theta} = \{\theta_1, ..., \theta_{n-1}\}$ where $C_k = \cos(\alpha_k)$ and $\tilde{C}_k = \cos(\theta_k)$. Then,

$S_k = \sin(\alpha_k)$ and $\tilde{S}_k = \sin(\theta_k)$, and the densities of the random angles α_k and θ_k are:

$$\begin{aligned} f_k^{\alpha} &= \frac{2}{B\left(\frac{\beta}{2}(n_1 - n + k), \frac{\beta}{2}(n_2 - n + k)\right)} \cos^{2a+\beta(k-1)+1}(\alpha_k) \sin^{2b+\beta(k-1)+1}(\alpha_k) \\ f_k^{\theta} &= \frac{2}{B\left(\frac{\beta}{2}k, \frac{\beta}{2}(n_1 + n_2 - 2n + k + 1)\right)} \cos^{\beta k}(\theta_k) \sin^{2a+2b+\beta(k-1)+3}(\theta_k) \end{aligned}$$

with a, b defined as in Theorem 3.1 and $B(x, y)$ the beta function. By independence, the joint density of $(\underline{\alpha}, \underline{\theta})$ is given by the product of the densities. For convenience we will denote

$$\tilde{Z}_{\beta,n} = \frac{2^{2n-1}}{\prod_{k=1}^{n} B\left(\frac{\beta}{2}(n_1 - n + k), \frac{\beta}{2}(n_2 - n + k)\right) \prod_{k=1}^{n-1} B\left(\frac{\beta}{2}k, \frac{\beta}{2}(n_1 + n_2 - 2n + k + 1)\right)}$$

This will be the necessary normalizing constant.

We will now map $(\underline{\alpha}, \underline{\theta})$ to the entries of M and from there to $(\underline{\lambda}, \underline{q})$ where the λ_i are the eigenvalues of MM^T and the q_i are the positive leading entries of the corresponding normalized eigenvectors. This second map will actually by the composition of several maps.

From the angles we map to the entries of the bidiagonal matrix:

Lemma A.2. *The Jacobian of the transform* $T : (\underline{\alpha}, \underline{\theta}) \to (\underline{x}, \underline{y})$, *where*

$$\begin{aligned} x_k &= \cos(\alpha_k) \sin(\theta_k), \quad k = 1, \cdots, n \\ y_k &= \sin(\alpha_{k+1}) \cos(\theta_k), \quad k = 1, \cdots, n-1. \end{aligned}$$

(for notational convenience we take $\theta_n = \pi/2$*) is given by*

$$J_T = \frac{\sin^2(\alpha_n)}{\sin(\alpha_1)} \prod_{k=1}^{n-1} \sin^2(\alpha_k) \sin^2(\theta_k).$$

Proof. The matrix of the partial derivatives with ordering $(\alpha_1, ..., \alpha_n, \theta_1, ..., \theta_{n-1}) \to (x_1, ..., x_n, y_1, ..., y_{n-1})$ has diagonal entries

$$-\sin\alpha_1 \sin\theta_1, ..., -\sin\alpha_{n-1} \sin\theta_{n-1}, -\sin\alpha_n, -\sin\alpha_2 \sin\theta_1, ..., -\sin\alpha_n \sin\theta_{n-1}.$$

Row and column reduction gives us that the determinant is given by the product of the diagonal, therefore

$$J_T = -\frac{\sin^2(\alpha_n)}{\sin(\alpha_1)} \prod_{k=1}^{n-1} \sin^2(\alpha_k)\sin^2(\theta_k) \qquad \square$$

We finish by mapping from the bidiagonal to the tridiagonal matrix, which in turn is mapped to by the eigenvalues. Denote by $q_1, ..., q_n$ the leading entries of the eigenvectors associated with the ordered eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ of the tridiagonal matrix normalized so that $q_i > 0$ and $\sum q_i^2 = 1$.

Lemma A.3 ([17], Lemmas 2.7, 2.9 and 2.11)**.** *For $\underline{x}, \underline{y}, \underline{a}, \underline{b}, \underline{\lambda}$ and $\underline{q}$ defined as above we have the following:*

1. *The Jacobian of the map $\psi : (\underline{x}, \underline{y}) \to (\underline{a}, \underline{b})$ can be written as*

$$J_\psi = 2^n x_1 \prod_{i=2}^{n} x_i^2.$$

2. *The Vandermonde determinant for the ordered eigvenvalues of a symmetric dridiagonal matrix with positive sub-diagonal $b = (b_{n-1}, ..., b_1)$ is given by*

$$\Delta(\underline{\lambda}) = \prod_{i<j}(\lambda_i - \lambda_j) = \frac{\prod_{i=1}^{n-1} b_i^i}{\prod_{i=1}^{n} q_i}.$$

3. *The Jacobian of the map $\phi : (\underline{a}, \underline{b}) \to (\underline{\lambda}, \underline{q})$ can be written as*

$$J_\phi = \frac{\prod_{i=1}^{n-1} b_i}{\prod_{i=1}^{n} q_i}.$$

Written in terms of our $\underline{x}, \underline{y}, \underline{a}$, and $\underline{b}$ with $Q_n = \prod_{i=1}^{n} q_i$ we get

$$J_\phi = \frac{1}{Q_n} \sin(\alpha_n)\cos(\alpha_n)\cos(\theta_{n-1}) \prod_{k=1}^{n-2} \sin(\alpha_{k+1})\cos(\alpha_{k+1})\sin(\theta_{k+1})\cos(\theta_k)$$

$$J_\psi = 2^n \cos(\alpha_1)\sin(\theta_1)\cos^2(\alpha_n) \prod_{k=2}^{n-1} \cos^2(\alpha_k)\sin^2(\theta_k).$$

The remainder of the proof of Theorem 3.1, is to show that the Jacobian of the transformation gives the desired result.

$$
\begin{aligned}
d(\underline{\lambda}, \underline{q}) &= \frac{J_\phi}{J_\psi \times J_T} d(\underline{\alpha}, \underline{\theta}) \\
&= \tilde{Z}_{\beta,n} \frac{2^{-n}}{Q_n^{1-\beta}} \left(\frac{1}{Q_n} \prod_{k=1}^{n} \cos^{k-1} \alpha_k \sin^{k-1} \alpha_k \prod_{k-1}^{n-1} \cos^k \theta_k \sin^{k-1} \theta_k \right)^\beta \\
&\qquad \times \quad \prod_{k=1}^{n} \cos^{2a} \alpha_k \sin^{2b} \alpha_k \times \prod_{k=1}^{n-1} \sin^{2a+2b} \theta_k.
\end{aligned}
$$

We substitute in the following pieces: first, notice that

$$
(\det M_{n,\beta})^2 = \prod_{i=1}^{n} \lambda_i = \prod_{k=1}^{n} \cos^2 \alpha_k \cdot \prod_{k=1}^{n-1} \sin^2 \theta_k.
$$

Then, applying Lemma A.1, the singular values of $M_{n,\beta}$ are the squares of the eigenvalues of the symmetric tridiagonal matrix L with zeroes in the main diagonal and

$$
C_n,\ S_n\tilde{C}_{n-1},\ C_{n-1}\tilde{S}_{n-1},\ \ldots,\ S_2\tilde{C}_1,\ C_1\tilde{S}_1,
$$

in the off-diagonal. We can check that $L + I$ can be written as AA^T where A is the bidiagonal matrix with

$$
1, S_n, \tilde{S}_{n-1}, S_{n-1}, \ldots, S_1, \qquad \text{and} \qquad C_n, \tilde{C}_{n-1}, C_{n-1}, \tilde{C}_{n-2}, \ldots, C_1
$$

in the diagonal and below the diagonal respectively. Using the characterization from Lemma A.1, we find that

$$
\det(L + I) = \prod_{\substack{k=1,\cdots,n \\ \epsilon=+,-}} 1 + \epsilon\sqrt{\lambda_k} = \prod_{k=1}^{n} (1 - \lambda_k) = (\det A)^2 = \prod_{k=1}^{n} \sin^2 \alpha_k \prod_{k=1}^{n-1} \sin^2 \theta_k.
$$

Remark A.4. *At this point one could again apply Lemma A.1 to A to find a $4n \times 4n$ matrix with zeros in the diagonal and*

$$
1, C_n, S_n, \tilde{C}_{n-1}, \tilde{S}_{n-1}, C_{n-1}, \ldots, C_1, S_1
$$

above and below the diagonal. This matrix will have eigenvalue $\pm\sqrt{1 \pm \sqrt{\lambda_k}}$.

Lastly recall ([17], Lemma 2.7) that we have

$$\Delta(\lambda) = \frac{1}{Q_n} \prod_{k=1}^{n-1} b_k^k = \frac{1}{Q_n} \prod_{k=1}^{n} \sin^{k-1}\alpha_k \cos^{k-1}\alpha_k \cdot \prod_{k=1}^{n-1} \cos^k\theta_k \cdot \prod_{k=1}^{n-1} \sin^{k-1}\theta_k.$$

Making all the appropriate substitutions this gives us that

$$\frac{J_\phi}{J_\psi \times J_T} d(\underline{\alpha}, \underline{\theta}) \quad = \quad \tilde{Z}_{\beta,n} \frac{2^{-n}}{Q_n^{1-\beta}} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \cdot \prod_{k=1}^{n} \lambda_k^a \cdot \prod_{k=1}^{n} (1-\lambda_k)^b,$$

From this we can see that the joint density function of $\underline{q}$ and $\underline{\lambda}$ separate. As in the Hermite and Laguerre cases we have that $\underline{q} \sim (\chi_\beta, ..., \chi_\beta)$ normalized to unit length [17]. This give us that for the unordered eigenvalues

$$f_{\beta,n,n_1,n_2}(\underline{\lambda}) = \frac{\tilde{Z}_{\beta,n}}{2^n n!} \frac{\left[\Gamma(\frac{\beta}{2})\right]^n}{\Gamma\left(\frac{\beta n}{2}\right)} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \cdot \prod_{k=1}^{n} \lambda_k^a \cdot \prod_{k=1}^{n} (1-\lambda_k)^b.$$

This completes the proof of Theorem 3.1. □

Remark A.5. *The normalizing constant on the final density can be written as*

$$Z_{\beta,n} = \left[\Gamma\left(\frac{\beta}{2}\right)\right]^n \prod_{k=1}^{n} \frac{\Gamma\left(\frac{\beta}{2}(n_1+n_2-n+k)\right)}{\Gamma\left(\frac{\beta}{2}(n_1-n+k)\right)\Gamma\left(\frac{\beta}{2}(n_2-n+k)\right)\Gamma\left(\frac{\beta}{2}k\right)}.$$

This gives an alternate derivation for the Selberg integral.

Appendix B

Properties of the LDP Rate Functions

B.1 All about $\mathcal{I}$

In this section we prove the needed estimates about the function $\mathcal{I}$.

Proposition B.1. *The function $\mathcal{I}(x)$ is strictly convex and continuous on $(0, \infty)$. It has an absolute minimum at $x = 1$ where it is equal to 0.*

Proof. K is strictly increasing on $(-\infty, 1)$ which shows that $\mathcal{I}(x)$ is well-defined (and differentiable) on $(0, \infty)$. By differentiating (4.3) and using the identities

$$K'(x) = \frac{E(x) - (1-x)K(x)}{2(1-x)x}, \qquad E'(x) = \frac{E(x) - K(x)}{2x}$$

we can compute that

$$\mathcal{I}'(x) = \frac{1}{x}\mathcal{I}(x) - \frac{1}{8x}K^{-1}\left(\frac{\pi}{2x}\right), \tag{B.1}$$

and

$$\mathcal{I}''(x) = \frac{\pi}{16x^3}\frac{1}{K'(K^{-1}(\frac{\pi}{2x}))}. \tag{B.2}$$

Observe that $K'(y) > 0$ for $y < 1$ which gives $\mathcal{I}''(x) > 0$ for $x > 0$ and the strict convexity of $\mathcal{I}$.

Using $K(0) = E(0) = \frac{\pi}{2}$ we get $\mathcal{I}(1) = \mathcal{I}'(1) = 0$ which (by the strict convexity) proves the second half of the proposition. $\square$

Proposition B.2. *We have* $\lim_{x\to 0^+} \mathcal{I}(x) = \frac{1}{8}$ *and* $\lim_{x\to 0^+} \mathcal{I}'(x) = -\frac{1}{2\pi}$. *There is a constant* $c_1 > 0$ *so that*

$$\mathcal{I}(x) \geq c_1(x-1)^2, \qquad \text{for all } x, \text{ and} \tag{B.3}$$

$$\frac{\mathcal{H}(-x)}{\sqrt{x}\log x}, \ \frac{\mathcal{I}(x)}{x^2\log^2 x} \text{ and } \frac{-K^{-1}(1/x)}{x^2\log^2 x} \text{ are bounded away from 0 and } \infty \text{ for } x > 2. \tag{B.4}$$

Proof. The following asymptotics can be readily derived from the definitions of elliptic integrals (or by the existing more sophisticated expansions c.f. [10],[31]). There is a constant $c > 0$ so that

$$\left|K(-a) - \frac{1}{2\sqrt{a}}\log(16a)\right| \leq \frac{c}{a^{3/2}}\log(a), \quad \left|E(-a) - \sqrt{a}\right| \leq \frac{c}{a^{1/2}}\log(a), \quad \text{for } a > 2. \tag{B.5}$$

From this it is easy to check that

$$\lim_{x\to\infty} \frac{\mathcal{H}(-x)}{\sqrt{x}\log x} = \frac{1}{2}, \qquad \lim_{x\to\infty} \frac{-K^{-1}(1/x)}{x^2\log^2 x} = 1, \quad \text{and} \qquad \lim_{x\to\infty} \frac{\mathcal{I}(x)}{x^2\log^2 x} = \frac{1}{2\pi^2}. \tag{B.6}$$

This gives (B.4). Note, that together with (B.1) this also gives

$$\lim_{x\to\infty} \frac{\mathcal{I}'(x)}{x\log^2 x} = \frac{1}{\pi^2}. \tag{B.7}$$

Using the functional identities

$$E(z) = \sqrt{1-z}\,E\left(\frac{z}{z-1}\right), \qquad K(z) = \frac{1}{\sqrt{1-z}}K\left(\frac{z}{z-1}\right), \qquad z \in (0,1)$$

the asymptotics of (B.5) can be transformed into

$$K(a) \sim -\tfrac{1}{2}\log(1-a), \qquad E(a) \sim 1, \qquad \text{as } a \to 1^{-1}$$

with explicit error bounds for $2/3 < a < 1$. From this we can obtain $\lim_{x\to 0^+} \mathcal{I}(x) = \frac{1}{8}$ and $\lim_{x\to 0^+} \mathcal{I}'(x) = -\frac{1}{2\pi}$. Using (B.4) with the continuity of $\mathcal{I}$ and the fact that $\mathcal{I}(1) = 0$ is an absolute minimum with $\mathcal{I}''(1) > 0$ gives (B.3). $\square$

The following two lemmas help to consolidate error terms that appear in the proofs Theorems 2.9 and 4.1.

Lemma B.3. *There exists an absolute constant c such that for any $t, q > 0$ we have*

$$|\mathcal{H}(a)| + |a|t/2 \le c(t+1)(\mathcal{I}(q)+1) \tag{B.8}$$

where $a = a(q) = K^{-1}(\pi/(2q))$.

Proof. Using (B.5) with the definition (4.3) we get that there is a constant c_2 so that

$$c_2^{-1}a(q) \le \mathcal{I}(q) \le c_2 a(q), \qquad \text{if } q > 2, \tag{B.9}$$

and the same bounds also give

$$|\mathcal{H}(a(q))| \le c_3\sqrt{|a(q)|}\log|a(q)| \tag{B.10}$$

for some constant c_3 in the same region. This shows the existence of a constant A with

$$|\mathcal{H}(a)| \le A\mathcal{I}(q), \qquad \text{and} \qquad a(q) \le A\mathcal{I}(q), \qquad \text{for } q > 2.$$

Since for $0 < q < 2$ both $a(q)$ and $\mathcal{H}(a(q))$ are bounded the lemma follows. $\square$

Lemma B.4. *For any $0 \le \varepsilon < 1/2$ there exists an absolute constant c, so that*

$$\mathcal{I}(x+\varepsilon) \le (1+\varepsilon)I(x) + c\varepsilon \tag{B.11}$$

Proof. Since $\mathcal{I}$ is convex, we have

$$\mathcal{I}(x+\varepsilon) \le \mathcal{I}(x) + \varepsilon \mathcal{I}'(x+\epsilon).$$

Since $\mathcal{I}(x)$ is decreasing on $[0, \pi/2]$, the bound (B.11) follows immediately for $x \in [0, \pi/2 - \varepsilon]$ with any $c \ge 0$. Using (B.1) we get

$$\mathcal{I}(x+\varepsilon) \le \mathcal{I}(x) + \frac{\varepsilon}{x+\varepsilon}\left(\mathcal{I}(x+\varepsilon) - \frac{1}{8}K^{-1}\left(\frac{\pi}{2(x+\varepsilon)}\right)\right). \tag{B.12}$$

From (B.6) it follows that there exists an $x_0 > 0$ such that

$$\frac{\varepsilon}{x+\varepsilon}\left(\mathcal{I}(x+\varepsilon) - \frac{1}{8}K^{-1}\left(\frac{\pi}{2(x+\varepsilon)}\right)\right) \le \varepsilon \mathcal{I}(x), \qquad \text{if } x \ge x_0$$

uniformly in $\varepsilon \in [0, 1/2]$. Therefore, for $x > x_0$ we have that $\mathcal{I}(x+\varepsilon) \le (1+\varepsilon)\mathcal{I}(x)$. We can assume $x_0 > \pi/2$. By choosing

$$c = \sup_{x \in [\pi/2, x_0+1/2]} \mathcal{I}'(x) = \mathcal{I}'(x_0 + 1/2)$$

we get $\mathcal{I}(x+\varepsilon) \le \mathcal{I}(x) + c\varepsilon$ on $[\pi/2 - \varepsilon, x_0]$ with any $0 \le \varepsilon < 1/2$ and the lemma follows. $\square$

B.2 Properties of I_{Sine}

In the final section of the appendix we describe the behavior of the function $I_{\text{Sine}}(\rho)$ near $\rho = \frac{1}{2\pi}$ and $\rho \to \infty$.

Proposition B.5. *The functions $\gamma(\nu)$ and $I_{Sine}(\rho)$ satisfy the following.*

1. *The function $\gamma(\nu)$ defined in (2.13) is continuous and strictly decreasing. It satisfies the differential equation $4x(1-x)\gamma''(x) = \gamma(x)$ on $(-\infty, 0)$ and on $(0,1)$ with boundary behavior $\lim_{x\to 0^\pm} \gamma(x) = \frac{1}{2\pi}$, $\gamma(1) = 0$ and $\lim_{x\to-\infty} \frac{\gamma(x)}{\sqrt{|x|}} = \frac{1}{4}$. These limits and the differential equation identify $\gamma(x)$ uniquely on $(-\infty, 0) \cup (0, 1]$.*

2. We have $I_{Sine}(0) = \frac{1}{64}$, $I_{Sine}(\frac{1}{2\pi}) = 0$, and $I''_{Sine}(x) > 0$ for $x \neq \frac{1}{2\pi}$. Moreover, we have the following limits:

$$\lim_{x\to 0} \frac{I_{Sine}(\frac{1}{2\pi}+x)}{\frac{x^2}{\log(1/|x|)}} = \frac{\pi^2}{4}, \qquad \text{and} \qquad \lim_{\rho\to\infty} \frac{I_{Sine}(\rho)}{\rho^2\log\rho} = \frac{1}{2}.$$

Proof. Recall the function $\gamma(\nu)$ given in (2.13). Using the asymptotics (B.6) proved in Proposition B.2 it is easy to see that $\gamma(\nu)$ is well-defined and positive in $(-\infty,0)\cup(0,1)$ with $\lim_{\nu\to 1^-}\gamma(\nu) = 0 = \gamma(1)$ and $\lim_{x\to-\infty}\frac{\gamma(x)}{\sqrt{|x|}} = \frac{1}{4}$. We also get that $\gamma(x)\mathcal{H}(x)$ blows up as $\nu \to 0^-$ or 0^+.

Differentiating (2.13) and using the definition (2.11) lead to

$$\mathcal{H}(x)\gamma'(x) = \frac{1}{8} + \mathcal{H}'(x)\gamma(x), \qquad \frac{\gamma''(x)}{\gamma(x)} = \frac{\mathcal{H}''(x)}{\mathcal{H}(x)} = \frac{1}{4x(1-x)}, \tag{B.13}$$

for $x \in (-\infty,0)\cup(0,1)$. We have $\mathcal{H}'(x) = -\frac{K(x)}{2} < 0$ for $x < 1$ and $\mathcal{H}(0) = 0$. Thus from the second identity we get that $\gamma'(\nu)$ is strictly decreasing in $(-\infty,0)$ and strictly increasing in $(0,1)$. From the asymptotics (B.6) of Proposition B.2 it is not hard to check that $\lim_{\nu\to-\infty}\gamma'(\nu) = 0$ and $\lim_{\nu\to 1^-}\gamma'(\nu) = -\frac{1}{8}$. This, together with the previous statement, proves that $\gamma(\nu)$ is decreasing on $(-\infty,0)$ and also on $(0,1]$.

Since $(\gamma(x)\mathcal{H}(x)^{-1})' = \frac{1}{8}\mathcal{H}(x)^{-2}$, L'Hospital's rule gives

$$\lim_{\nu\to 0}\gamma(\nu) = \lim_{\nu\to 0}\frac{\gamma(\nu)\mathcal{H}(\nu)^{-1}}{\mathcal{H}(\nu)^{-1}} = -\tfrac{1}{8}H'(0) = \frac{1}{2\pi} = \gamma(0).$$

Then from (B.13) it follows that

$$\lim_{\nu\to 0}\gamma''(\nu)\nu = \frac{1}{8\pi}, \qquad \lim_{\nu\to 0}\frac{\gamma'(\nu)}{\log|\nu|} = \frac{1}{8\pi},$$

and also that

$$\lim_{x\to 0}\frac{\gamma^{(-1)}(\frac{1}{2\pi}+x)}{8\pi\frac{x}{\log|x|}} = 1. \tag{B.14}$$

This shows that $\gamma(\nu)$ is continuous and strictly decreasing on $(-\infty, 1]$. We have $\gamma^{(-1)}(1) = 0$ and $I_{\text{Sine}}(0) = \frac{1}{64}$.

Note, the fact that $\gamma(x)$ solves $4x(1-x)\gamma''(x) = \gamma(x)$ on $(-\infty, 0) \cup (0, 1)$ and has the proven asymptotics at $-\infty, 0$ and 1 uniquely identifies it. The equation $4x(1-x)y''(x) = y(x)$ has two linearly independent solutions on both $(-\infty, 0)$ and $(0, 1)$. The function $\mathcal{H}(x)$ also solves the equation (on both intervals), but with $\mathcal{H}(0) = 0$, $\lim_{x \to 1^-} \mathcal{H}(x) = -1$ and $\lim_{x \to -\infty} \frac{\mathcal{H}(x)}{\sqrt{|x|}\log|x|} = \frac{1}{2}$. This shows that any solution on $(-\infty, 0)$ or $(0, 1)$ can be expressed as $c_1\gamma + c_2\mathcal{H}$ with some constants c_1, c_2, and the values of the constants are determined by the behavior of the solution at the end of the interval.

Using (B.13) together with (2.15) we can also compute that

$$\begin{aligned} I'_{\text{Sine}}(\rho) &= \frac{1}{8}\left[\frac{1}{8\gamma'(\nu)} + \mathcal{H}(\nu) + \frac{\gamma(\nu)\mathcal{H}'(\nu)}{\gamma'(\nu)}\right] = \frac{1}{4}\mathcal{H}(\nu), \text{ and} \\ I''_{\text{Sine}}(\rho) &= \frac{1}{4}\frac{\mathcal{H}'(\nu)}{\gamma'(\nu)} = -\frac{1}{8}\frac{K(\nu)}{\gamma'(\nu)}, \end{aligned}$$

where ν is short for $\gamma^{-1}(\rho)$. From this $I_{\text{Sine}}(\frac{1}{2\pi}) = 0$ follows, together with $I_{\text{Sine}}(x) > 0$ for $x \neq \frac{1}{2\pi}$. The asymptotics of $I_{\text{Sine}}(\frac{1}{2\pi} + x)$ as $x \to 0$ can be obtained from the definition (2.15), the asymptotics (B.14), and the fact that $\mathcal{H}(0) = 0, \mathcal{H}'(0) = -\frac{\pi}{4}$.

Lastly we can look at the asymptotics of $I_{\text{Sine}}(\rho)$ as $\rho \to \infty$. Recalling again (B.6) we get

$$\mathcal{H}(-x) \sim \tfrac{1}{2}\sqrt{x}\log x, \qquad \gamma(-x) \sim \frac{\sqrt{x}}{4}, \qquad \gamma^{(-1)}(x) \sim 16x^2 \qquad \text{as } x \to \infty,$$

from which $I_{\text{Sine}}(\rho) \sim \frac{1}{2}\rho^2 \log \rho$ follows for $\rho \to \infty$. $\square$

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of mathematical function with formulas, graphs, and mathematical tables.* Wiley, 1984.

[2] G. Akemann, J. Baik, and P. D. Francesco, editors. *The Oxford Handbook of Random Matrix Theory.* Oxford University Press, Oxford, 2011.

[3] G. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices.* Cambridge Studies in Advanced Mathematics, 118. Cambridge University Press, Cambridge, 2009.

[4] E. L. Basor, C. A. Tracy, and H. Widom. Asymptotics of level-spacing distributions for random matrices. *Phys. Rev. Lett.*, 69(1):5–8, 1992.

[5] G. Ben Arous and A. Guionnet. Large deviations for Wigner's law and Voiculescu's non-commutative entropy. *Probab. Theory Related Fields*, 108(4):517–542, 1997.

[6] A. Bloemendal and B. Virág. Limits of spiked random matrices. *Probab. Theory Related Fields*, 156(3-4):795–825, 2013.

[7] P. Bourgade, L. Erdős, and H.-T. Yau. Bulk universality of general β-ensembles with non-convex potential. *J. Math. Phys.*, 53:095221, September 2012.

[8] P. Bourgade, L. Erdős, and H.-T. Yau. Edge universality of beta ensembles. *http://arxiv.org/abs/1306.5728*, June 2013.

[9] G. Buttazzo, M. Giaquinta, and S. Hildebrant. *One-dimensional Variational Problems*. Oxford lecture series in mathematics and its applications. Clarendon Press, Oxford, 1998.

[10] B. C. Carlson and J. L. Gustafson. Asymptotic expansion of the first elliptic integral. *SIAM J. Math Anal.*, 16(5):1072–1092, September 1985.

[11] B. Collins. Product of random projections, jacobi ensembles and universality problems arising from free probability. *Probab. Theory Related Fields*, 133:315–344, 2005.

[12] P. Deift, A. Its, I. Krasovsky, and X. Zhou. The Widom-Dyson constant for the gap probability in random matrix theory. *J. Comput. Appl. Math.*, 202(1):26–47, 2007.

[13] P. Deift, A. Its, and X. Zhou. A Riemann-Hilbert approach to asymptotic problems arising in the theory of random matrices and also in the theory of integrable statistical mechanics. *Ann. Math.*, 146:149–235, 1997.

[14] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38. Springer, 2010.

[15] H. Dette and J. Nagel. Some asymptotic properties of the spectrum of the jacobi ensemble. *SIAM J. Math. Anal.*, 41(4), 2009.

[16] I. Dumitriu. Smallest eigenvalue distributions for two classes of β-jacobi ensembles. *J. Math. Phys.*, 53(10), 2012.

[17] I. Dumitriu and A. Edelman. Matrix models for beta ensembles. *J. Math. Phys.*, 43(11):5830–5847, 2002.

[18] I. Dumitriu and P. J. Forrester. Tridiagonal realization of the antisymmetric Gaussian β-ensemble. *J. Math. Phys.*, 51(9):093302, 25, 2010.

[19] I. Dumitriu and P. Koev. Distribution of the extreme eigenvalues of the beta-jacobi random matrix model. *SIAM J. Matrix Anal. Appl.*, 30(1):1–6, 2008.

[20] F. J. Dyson. The Coulumb fluid and the fifth Painlevé transcendent. In *Chen Ning Yang*, pages 317–341. Int. Press, Cambridge, MA, 1995.

[21] A. Edelman and B. D. Sutton. From random matrices to stochastic operators, 2007. math-ph/0607038.

[22] T. Ehrhardt. Dyson's constant in the asymptotics of the Fredholm determinant of the sine kernel. *Comm. Math. Phys.*, 262(2):317–341, 2006.

[23] L. Erdős. Universality for random matrices and log-gases. *http://arxiv.org/abs/1212.0839*, Dec 2012.

[24] L. Erdős, A. Knowles, H.-T. Yau, and J. Yin. The local semicircle law for a general class of random matrices. *Electron. J. Probab.*, 18(59):1–58, May 2013.

[25] L. Erdős, H.-T. Yau, and J. Yin. Bulk universality for generalized wigner matrices. *Probab. Theory Related Fields*, 154(1-2):341–407, October 2012.

[26] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[27] M. M. Fogler and B. I. Shklovskii. Probability of an eigenvalue number fluctuation in an interval of a random matrix spectrum. *Phys. Rev. Lett.*, 74(17):3312–3315, Apr 1995.

[28] P. J. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010.

[29] P. J. Forrester and N. S. Witte. Asymptotic forms for hard and soft edge general î conditional gap probabilities. *Nuclear Phys. B*, 849(3):321–340, 2012.

[30] S. Graversen and A. Shiryaev. An extension of p. lévy's distributional. *Bernoulli*, 6(4):615–620, 2000.

[31] J. L. Gustafson. *Asymptotic formulas for elliptic integrals.* PhD thesis, Iowa State University, Ames, Iowa, 1982.

[32] D. Holcomb and G. R. M. Flores. Edge scaling of the β-jacobi ensemble. *J Stat Phy*, 149:1136–1160, Dec 2012.

[33] D. Holcomb and B. Valkó. Large deviations for the sine_β and sch_τ processes. *http://arxiv.org/abs/1311.2981*, 2013.

[34] J. Hough, M. Krishnapur, Y. Peres, and B. Virág. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes.* University Lecture Series, 51. American Math. Soc., Providence, RI, 2009.

[35] S. Jacquot and B. Valkó. Bulk scaling limit of the Laguerre ensemble. *Electron. J. Probab.*, 16:no. 11, 314–346, 2011.

[36] T. Jiang. Limit theorems for Beta-Jacobi ensembles. *Bernoulli*, 19(3):1028–1046, 2013.

[37] I. Johnstone. Multivariate analysis and jacobi ensembles: largest eigenvalue, tracy-widom limits and rates of convergence. *Ann. Statist.*, 36(6):2638–2716, 2008.

[38] I. Karatzas and J. Ruf. Distribution of the time to explosion for one-dimensional diffusions. *http://arxiv.org/abs/1303.5899*, 03 2013.

[39] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[40] R. Killip and I. Nenciu. CMV: the unitary analogue of Jacobi matrices. *Comm. Pure Appl. Math.*, 60(8):1148–1188, 2007.

[41] I. V. Krasovsky. Gap probability in the spectrum of random matrices and asymptotics of polynomials orthogonal on an arc of the unit circle. *Int. Math. Res. Not.*, (25):1249–1272, 2004.

[42] M. Krishnapur, B. Rider, and B. Virág. Universality of the stochastic airy operator. *http://arxiv.org/abs/1306.4832*, June 2013.

[43] E. Kritchevski, B. Valkó, and B. Virág. The scaling limit of the critical one-dimensional random Schrödinger operator. *Comm. Math. Phys.*, 314(3):775–806, 2012.

[44] M. L. Mehta. *Random matrices*, volume 142 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, third edition, 2004.

[45] T. Nagao and P. Forrester. Asymptotic correlations at the spectrum edge of random matrices. *Nuclear Phys. B*, 435(3):401–420, 1995.

[46] T. Nagao and M. Wadati. Eigenvalue distribution of random matrices at the spectrum edge. *J. Phys. Soc. Japan*, 62:3845–3856, 1993.

[47] J. Ramírez, B. Rider, and O. Zeitouni. Hard edge tail asymptotics. *Electron. Commun. Probab.*, 16, 2011.

[48] J. A. Ramírez and B. Rider. Diffusion at the random matrix hard edge. *Comm. Math. Phys.*, 288(3):887–906, 2009.

[49] J. A. Ramírez, B. Rider, and B. Virág. Beta ensembles, stochastic Airy spectrum, and a diffusion. *J. Amer. Math. Soc.*, 24(4):919–944, 2011.

[50] M. M. Rao and Z. D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1991.

[51] B. D. Sutton. *The stochastic operator approach to random matrix theory*. ProQuest LLC, Ann Arbor, MI, 2005. Thesis (Ph.D.)–Massachusetts Institute of Technology.

[52] T. Tao and V. Vu. Universality of local spectral statistics of non-hermitian matrices. *Bull. Amer. Math. Soc.*, 49:377–414, Jan 2012.

[53] C. A. Tracy and H. Widom. Introduction to random matrices. In *Geometric and quantum aspects of integrable systems (Scheveningen, 1992)*, volume 424 of *Lecture Notes in Phys.*, pages 103–130. Springer, Berlin, 1993.

[54] B. Valkó and B. Virág. Continuum limits of random matrices and the Brownian carousel. *Inventiones Math.*, 177:463–508, 2009.

[55] B. Valkó and B. Virág. Large gaps between random eigenvalues. *Ann. Probab.*, 38(3):1263–1279, 2010.

[56] H. Widom. The asymptotics of a continuous analogue of orthogonal polynomials. *J. Approx. Theory*, 77:51–64, 1996.

[57] E. P. Wigner. *Gatlinberg Conference on Neutron Physics, Oak Ridge National Laboratory Report*, ORNL 2309:59, 1957.

[58] Wishart. The generalized product moment distribution in samples from a normal multivariate population. *Biometrica*, 20A:32–152, 1928.