

GAUSSIAN FLUCTUATIONS FOR PRODUCTS OF RANDOM MATRICES

VADIM GORIN AND YI SUN

ABSTRACT. We study global fluctuations for singular values of M -fold products of several right-unitarily invariant $N \times N$ random matrix ensembles. As $N \rightarrow \infty$, we show the fluctuations of their height functions converge to an explicit Gaussian field, which is log-correlated for M fixed and has a white noise component for $M \rightarrow \infty$ jointly with N . Our technique centers on the study of the multivariate Bessel generating functions of these spectral measures, for which we prove a central limit theorem for global fluctuations via certain conditions on the generating functions. We apply our approach to a number of ensembles, including square roots of Wishart, Jacobi, and unitarily invariant positive definite matrices with fixed spectrum, using a detailed asymptotic analysis of multivariate Bessel functions to verify the necessary conditions.

CONTENTS

1. Introduction	2
1.1. Overview of the results	2
1.2. Comparison with fluctuations for sums of random matrices	3
1.3. Central limit theorems for multivariate Bessel generating functions	4
1.4. Asymptotics of multivariate Bessel generating functions for matrix products	6
1.5. Multivariate Bessel generating functions and the Cholesky decomposition	6
1.6. Relation to the literature	7
1.7. Organization of the paper	9
Acknowledgments	9
2. Central limit theorem via multivariate Bessel generating functions	9
2.1. Multivariate Bessel generating functions	9
2.2. Statement of the results	11
2.3. The symmetrization procedure on analytic functions	12
2.4. A graphical calculus	13
2.5. Proof of Theorem 2.5	16
2.6. Proof of Theorem 2.6	19
2.7. Central limit theorems for growing measures	22
3. Asymptotics of multivariate Bessel functions	24
3.1. Statement of the results	24
3.2. Multivariate Bessel functions and hook Schur functions	26
3.3. An integral formula for multivariate Bessel functions	28
3.4. Single variable asymptotics	30
3.5. Proof of Theorem 3.4	40
4. Products of random matrices	41
4.1. Multivariate Bessel generating functions for matrix products	41
4.2. LLN-appropriate and CLT-appropriate measures	42
4.3. Products of finitely many random matrices	45
4.4. Lyapunov exponents for matrices with fixed spectrum	53
4.5. Lyapunov exponents for Jacobi and Ginibre	57
5. A 2-D Gaussian field from products of random matrices	60
5.1. Statement of the result	60
5.2. Multilevel LLN and CLT via multivariate Bessel generating functions	60
5.3. Proof of Theorem 5.1	62
Appendix A. Fluctuations of sums of random matrices	63

Appendix B. Multivariate Bessel generating functions and the Cholesky decomposition	64
References	66

1. INTRODUCTION

1.1. Overview of the results. Let $Y_N^1, \dots, Y_N^M$ be i.i.d. $N \times N$ random matrices which are right-unitarily invariant, and let $Y_N := Y_N^1 \cdots Y_N^M$ be their product. The (squared) singular values $\mu_1^N \geq \dots \geq \mu_N^N > 0$ of Y_N occur classically in the study of ergodic theory of non-commutative random walks, and the corresponding Lyapunov exponents have a limit as $M \rightarrow \infty$ by Oseledec's multiplicative ergodic theorem (see [Ose68, Rag79]). These Lyapunov exponents for various ensembles of matrices have been the object of extensive study in the dynamical systems literature, beginning with the pioneering result of Furstenberg-Kesten in [FK60] for matrices with positive entries. In an applied context, singular values for similar models have appeared in the study of disordered systems in statistical physics as described in [CPV93], in the study of polymers as in [CMR17], and in the study of dynamical isometry for deep neural networks as the Jacobians of randomly initialized networks (see [SMG14, PSG17, PSG18, CPS18, XBSD⁺18, TWJ⁺18, LQ18]).

The goal of this work is to study the global fluctuations of the squared singular values μ^N . The distribution of μ^N depends only on the distribution of $X_N^i := (Y_N^i)^* Y_N^i$, and we study it for a variety of different distributions for X_N^i , including Wishart matrices, Jacobi matrices, and unitarily invariant positive definite matrices with fixed spectrum. In each of these cases, we study the normalized log-spectrum

$$\lambda_i^N := \frac{1}{M} \log \mu_i^N$$

via the height function

$$\mathcal{H}_N(t) := \#\{\lambda_i^N \leq t\}.$$

We study limit shapes and fluctuations for the height function in two limit regimes, one where $N \rightarrow \infty$ with M fixed, and one where $N, M \rightarrow \infty$ simultaneously. Note that in the main text we will use a slightly different notation for the log-spectrum and height function for M fixed which removes the normalization by M , but we keep this uniform notation in the introduction for clarity of exposition.

When M is fixed, results from free probability of Voiculescu and Nica-Speicher in [Voi87, NS97] imply that the empirical measure $d\lambda^N$ of λ_i^N converges to a deterministic measure $d\lambda^\infty$, which implies that $\mathcal{H}_N(t)$ concentrates around a deterministic limit shape, a result we refer to as a law of large numbers. In Theorems 4.8, 4.12, and 4.13 and Corollary 4.9, we show that the fluctuations of the height function around its mean converge to explicit Gaussian fields. In Corollary 4.11, we show that these fields are log-correlated under a technical condition on the smoothness of the limit shape. Namely, this means that for polynomials f, g , we have

$$\lim_{N \rightarrow \infty} \text{Cov} \left(\int (\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) f(t) dt, \int (\mathcal{H}_N(s) - \mathbb{E}[\mathcal{H}_N(s)]) g(s) ds \right) = \int \int K(t, s) f(t) g(s) dt ds$$

for an explicit covariance kernel $K(t, s)$ which satisfies

$$K(t, s) = -\frac{1}{2\pi^2} \log |t - s| + O(1)$$

for $|t - s| \rightarrow 0$.

When $M, N \rightarrow \infty$ simultaneously, the λ_i^N are the Lyapunov exponents studied by Newman and Isopi-Newman in [New86a, New86b, IN92] and mentioned by Deift in [Dei17]. Again, their empirical measure converges to the deterministic measure

$$-\frac{e^{-z}}{S'_\mu(S_\mu^{-1}(e^{-z}))} \mathbf{1}_{[-\log S_\mu(-1), -\log S_\mu(0)]} dz,$$

where S_μ is the limiting S -transform of the spectral measure of $(Y_N^i)^* Y_N^i$; we give a more detailed treatment in Sections 4.4 and 4.5. This yields concentration of the height function around a deterministic limit shape. In Theorems 4.14, 4.16, and 4.17, we show that the rescaled fluctuations of the height function around its mean

$$M^{1/2} \left(\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)] \right)$$

again converge to explicit Gaussian fields. However, in these cases, we find that the field has a white noise component, meaning that for $|t - s| \rightarrow 0$ the covariance satisfies

$$K(t, s) = \delta(t - s) + O(1),$$

where δ denotes the Dirac delta function. In particular, as $M \rightarrow \infty$, we see a transition from a log-correlated Gaussian field to a Gaussian field with a white noise component. For this result, the relative growth rates of N and M are not important; in particular, we observe the same white noise component both for M growing to infinity much faster than N and much slower than N . This behavior is very different from that seen in recent work of Akemann-Burda-Kieburg and Liu-Wang-Wang in [ABK18, LWW18], where the local limit for the singular values of the products of Ginibre matrices depends on the ratio M/N .

1.2. Comparison with fluctuations for sums of random matrices. We may naturally compare our results to analogous results for sums of random matrices. For $i = 1, \dots, M$, let $A_N^i = U_i A U_i^*$ be $N \times N$ matrices with U_i i.i.d. Haar unitary and A deterministic diagonal with spectrum μ^N . Define

$$X_{N,M}^{\text{add}} := A_N^1 + \dots + A_N^M$$

and the height function of the spectrum $\lambda^{N,M}$ of $X_{N,M}^{\text{add}}$ by $\mathcal{H}_{N,M}(t) := \#\{\lambda_i^{N,M} \leq t\}$. If M is finite and fixed and the empirical measures of μ^N converge to a measure $d\mu$, results from free probability show that $\mathcal{H}_{N,M}(t)$ converges as $N \rightarrow \infty$ to a deterministic limit shape. It is further known that the random variables

$$\{\mathcal{H}_{N,M}(t) - \mathbb{E}[\mathcal{H}_{N,M}(t)]\}_{t \in \mathbb{R}}$$

converge to an explicit log-correlated Gaussian field given by Collins-Mingo-Śniady-Speicher in [CMSS07] and Pastur-Vasilchuk in [PV07].

On the other hand, if we take $M \rightarrow \infty$, then we have the convergence

$$\frac{1}{M} X_{N,M}^{\text{add}} \rightarrow \frac{1}{N} \left(\sum_{i=1}^N \mu_i^N \right) \cdot \text{Id}$$

of the average $\frac{1}{M} X_{N,M}^{\text{add}}$ of the matrices A_N^i to a deterministic matrix. In contrast to the multiplicative case, this averaging retains only the first moment of the empirical measure of the original spectrum. Furthermore, by the ordinary central limit theorem, as $M \rightarrow \infty$ with N fixed, the random matrix

$$\frac{1}{\sqrt{M}} \left(X_{N,M}^{\text{add}} - \mathbb{E}[X_{N,M}^{\text{add}}] \right)$$

converges in distribution to a Gaussian random matrix which we verify in Appendix A to have distribution

$$\sqrt{\frac{N}{(N-1)(N+1)}} \cdot \left[\frac{1}{N} \sum_{i=1}^N (\mu_i^N)^2 - \frac{1}{N^2} \left(\sum_{i=1}^N \mu_i^N \right)^2 \right] \cdot \text{GUE}_{N, \text{Tr}=0},$$

where $\text{GUE}_{N, \text{Tr}=0} \stackrel{d}{=} X - \text{Tr}(X) \cdot \text{Id}_N$ for $X \stackrel{d}{=} \text{GUE}_N$ being an $N \times N$ matrix drawn from the Gaussian Unitary Ensemble. The results of Johansson in [Joh98] on the GUE then suggest that as $N \rightarrow \infty$ slowly compared with M , fluctuations of the corresponding height function remain a log-correlated Gaussian field.

One may ask why white noise shows up for $M \rightarrow \infty$ in the multiplicative setting but not in the additive setting. One intuitive explanation comes from comparing the decompositions

$$X_{N,M}^{\text{add}} = \mathbb{E}[X_{N,M}^{\text{add}}] + \left(X_{N,M}^{\text{add}} - \mathbb{E}[X_{N,M}^{\text{add}}] \right)$$

$$\log X_{N,M}^{\text{mult}} = \mathbb{E}[\log X_{N,M}^{\text{mult}}] + \left(\log X_{N,M}^{\text{mult}} - \mathbb{E}[\log X_{N,M}^{\text{mult}}] \right)$$

in the additive and multiplicative settings, where $X_{N,M}^{\text{mult}} := (Y_N^1 \cdots Y_N^M)^* (Y_N^1 \cdots Y_N^M)$.

In the additive setting, note that $\mathbb{E}[X_{N,M}^{\text{add}}]$ is a constant multiple of the identity, and therefore the $[tN]^{\text{th}}$ eigenvalue of $X_{N,M}^{\text{add}}$ admits the approximate representation

$$\lambda_{[tN]}^{\text{add}} \sim MC_1 + \sqrt{M}C_2 \cdot \gamma_{[tN]}$$

for constants C_1 and C_2 , where γ_i is the i^{th} eigenvalue of $\frac{1}{\sqrt{N}} \text{GUE}_{N, \text{Tr}=0}$. In particular, all fluctuations of the height function come from fluctuations of the height function for $X_{N,M}^{\text{add}} - \mathbb{E}[X_{N,M}^{\text{add}}]$.

In contrast, in the multiplicative setting, $\mathbb{E}[\log X_{N,M}^{\text{mult}}]$ has non-trivial spectrum, and the ordinary central limit theorem does not directly apply to $\log X_{N,M}^{\text{mult}} - \mathbb{E}[\log X_{N,M}^{\text{mult}}]$. Together, these two effects imply the analysis of the additive case does not apply in the multiplicative setting. It would be interesting to find other geometric qualitative arguments explaining the appearance of the white noise component in the products without relying on the exact formulas that we use. We will not pursue this direction here, only mentioning that the viewpoint of Reddy in [Red16] for fixed N seems relevant for such a potential development.

Another heuristic explanation for the appearance of the white noise and for the difference between additive and multiplicative cases can be obtained by comparing our model to Dyson Brownian Motion (DBM) started from a deterministic initial condition¹. In this approach, we should treat the number of terms or factors as inverse time, $t = M^{-1}$. For addition, such an identification is possible because of the invariance of the Brownian motion with respect to the time inversion $B_{1/t} \stackrel{d}{=} t^{-1}B_t$, $t \geq 0$. For multiplication, this is more speculative. However, for the special case when the factors are drawn from the square Ginibre ensemble (i.e. with i.i.d. Gaussian entries), one can compare the contour integral formulas for the correlation kernel for products given by Akemann-Kieburg-Wei and Kuijlaars-Zhang in [AKW13, KZ14] with the formulas given by Brézin-Hikami in [BH96, BH97] for the $\beta = 2$ Dyson Brownian Motion started from deterministic initial condition and notice their striking similarity upon making the identification $t = M^{-1}$. In addition, in this case Akemann-Burda-Kieburg found in [ABK18] that when $N, M \rightarrow \infty$ with N/M limiting to a value in $(0, \infty)$, the local statistics for products of square Ginibre matrices coincide with those of Dyson Brownian motion started from an evenly spaced initial condition as studied by Johansson in [Joh04].

For general products of matrices, this suggests that we should start the DBM from the initial condition given by the Lyapunov exponents. The transition between white noise (small t , corresponding to $M \rightarrow \infty$) and log-correlated (finite t , corresponding to finite M) statistics was studied in detail for $\beta = 2$ DBM by Duits-Johansson in [DJ18], for general β DBM by Huang-Landon in [HL16], and for a finite-temperature version of the GUE ensemble by Johansson-Lambert in [JL18]. In contrast, for the sums, the initial condition for the DBM is a diagonal matrix with all equal eigenvalues; hence, the global fluctuations are the same as those of the GUE, and we see log-correlated structure for the local statistics at all positive times. It would be extremely interesting to find a conceptual explanation for this analogy between products of matrices and Dyson Brownian Motion.

1.3. Central limit theorems for multivariate Bessel generating functions. Our technique is based on the study of multivariate Bessel generating functions for log-spectral measures, which are continuous versions of the Schur generating function defined and studied by Bufetov-Gorin in [BG15b, BG18a, BG18b]. Recall that for sets of variables $a = (a_1, \dots, a_N)$ and $b = (b_1, \dots, b_N)$, the multivariate Bessel function is defined by

$$\mathcal{B}(a, b) := \Delta(\rho) \frac{\det(e^{a_i b_j})_{i,j=1}^N}{\Delta(a)\Delta(b)},$$

where $\rho = (N-1, \dots, 0)$ and $\Delta(a) := \prod_{1 \leq i < j \leq N} (a_i - a_j)$ denotes the Vandermonde determinant. For an N -tuple $\chi_N = (\chi_{N,1} \geq \dots \geq \chi_{N,N})$, the multivariate Bessel generating function of a measure $d\mu_N(x)$ on N -tuples $x = (x_1 \geq \dots \geq x_N)$ with respect to χ_N is defined by

$$\phi_{\chi,N}(s) := \int \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi_N, x)} d\mu_N(x).$$

We say that the measure $d\mu_N$ is χ -smooth if the defining integral of $\phi_{\chi,N}(s)$ converges absolutely and uniformly on a neighborhood of $[\chi_{N,N}, \chi_{1,N}]^N$. For a sequence of N -tuples $\chi_N \in \{x \in \mathbb{R}^N \mid x_1 \geq \dots \geq x_N\}$ for which $\frac{1}{N} \sum_{i=1}^N \delta_{\chi_{N,i}/N} \rightarrow d\chi$ for a measure $d\chi$, we define a family of CLT-appropriate measures for χ_N by the following condition, which governs the behavior of their multivariate Bessel generating functions near $s = \chi_N$. For a subset $I \subset \{1, \dots, N\}$, define

$$\phi_{\chi,N}^I(s) := \phi_{\chi,N}(s)|_{s_j = \chi_{N,j}, j \notin I},$$

where we note that $\phi_{\chi,N}^I(s)$ is a function of $(s_i)_{i \in I}$ alone. For a function $f(s)$ on $\mathbb{C}^N$ and some $r \in \mathbb{C}^N$, we will adopt the notation that $\partial_{r_i} f(r) = \frac{\partial}{\partial s_i} f(s) \Big|_{s=r}$.

¹We thank Maurice Duits for communicating this point of view to us.

Definition 2.3. We say that the measures $d\mu_N$ are LLN-appropriate for χ_N if they are χ_N -smooth and there exist a compact set $V_\chi \subset \mathbb{R}$ and a holomorphic function Ψ on an open complex neighborhood U of V_χ such that $\chi_{N,i}/N \in V_\chi$ for all N and for each fixed integer $k > 0$, uniformly in I with $|I| = k$, $i \in I$, and $r^I \in U^{|I|}$, we have

$$\lim_{N \rightarrow \infty} \left| \frac{1}{N} \partial_{r_i} [\log \phi_{\chi,N}^I(rN)] - \Psi(r_i) \right| = 0.$$

Definition 2.4. We say that measures $d\mu_N$ are CLT-appropriate for χ_N if they are LLN-appropriate for χ_N with respect to some function Ψ and sets $V_\chi \subset \mathbb{R}$ with neighborhood U and there exists a holomorphic function Λ on U^2 such that for each fixed integer $k > 0$, uniformly in I with $|I| = k$, $r^I \in U^{|I|}$, and distinct $i, j \in I$, we have

$$\lim_{N \rightarrow \infty} \left| \partial_{r_i} \partial_{r_j} [\log \phi_{\chi,N}^I(rN)] - \Lambda(r_i, r_j) \right| = 0.$$

Our first technical tools are Theorems 2.5 and 2.6 giving a law of large numbers and central limit theorem for the moments $p_k := \int x^k d\mu$ of CLT-appropriate measures. Define the Cauchy transform of $d\chi$ by

$$\Xi(u) := \int \frac{1}{u-x} d\chi(x).$$

Theorem 2.5 (Law of large numbers). If the measures $d\mu_N$ are LLN-appropriate for χ_N , then we have the following convergence in probability

$$\lim_{N \rightarrow \infty} \frac{1}{N} p_k(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(x)] = \mathfrak{p}_k := \frac{1}{k+1} \oint \left(\Xi(u) + \Psi(u) \right)^{k+1} \frac{du}{2\pi i},$$

where the u -contour encloses V_χ and lies within U . In addition, the random measures $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ with x distributed according to $d\mu_N$ converge in probability to a deterministic compactly supported measure $d\mu$ with $\int x^k d\mu(x) = \mathfrak{p}_k$.

Theorem 2.6. If the measures $d\mu_N$ are CLT-appropriate for χ_N , then the collection of random variables

$$\{p_k(x) - \mathbb{E}[p_k(x)]\}_{k \in \mathbb{N}}$$

converges in the sense of moments to the Gaussian vector with zero mean and covariance

$$\lim_{N \rightarrow \infty} \text{Cov}(p_k(x), p_l(x)) = \mathfrak{Cov}_{k,l} := \oint \oint \left(\Xi(u) + \Psi(u) \right)^l \left(\Xi(w) + \Psi(w) \right)^k \left(\frac{1}{(u-w)^2} + \Lambda(w, u) \right) \frac{dw}{2\pi i} \frac{du}{2\pi i},$$

where the u and w -contours enclose V_χ and lie within U , and the u -contour is contained inside the w -contour.

Theorems 2.5 and 2.6 are parallel to results of [BG18a] for Schur generating functions. However, their proofs are significantly different. In [BG18a], the Schur functions are normalized by their evaluation at $\chi_N = (1, \dots, 1)$, while we allow general N -tuples χ_N ; in particular, for matrix products, we will choose $\chi_N = \rho$. Due to this feature, adapting the proofs requires several new ideas. In particular, our analysis hinges on the asymptotics of derivatives of $\log \phi_{\chi,N}(s)$ in at most two variables s_i, s_j evaluated at the point $s = \chi_N$. In our setting, such derivatives acquire a dependence on the choice of indices i, j from the fact that $\chi_{N,i}$ may be non-constant, which did not exist in the setting of [BG18a]. Handling this in the asymptotics gives rise to the integrals along contours surrounding V_χ in Theorem 2.6.

We derive a similar statement in Theorem 2.15 for the case where the form of the multivariate Bessel function changes with N . Let $M, N \rightarrow \infty$ simultaneously, and suppose there exists a sequence of measures $d\lambda'_N$ with multivariate Bessel generating function $\phi_{\chi,N}(s)^M$. Let $d\lambda_N$ be the pushforward of $d\lambda'_N$ under the map $\lambda' \mapsto \frac{1}{M} \lambda'$. We obtain a similar law of large numbers and central limit theorem for the moments of $d\lambda_N$, though we require a different scaling.

Theorem 2.14. If $d\mu_N$ is LLN-appropriate for χ_N , then if x is distributed according to $d\lambda_N$, in probability we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} p_k(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(x)] = \mathfrak{p}'_k := \oint \Xi(u) \Psi(u)^k \frac{du}{2\pi i},$$

where the u -contour encloses V_χ and lies within U . In addition, the random measures $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ with x distributed according to $d\lambda_N$ converge in probability to a deterministic compactly supported measure $d\lambda$ with $\int x^k d\lambda(x) = \mathbf{p}'_k$.

Theorem 2.15. If the measures $d\mu_N$ are CLT-appropriate for χ_N , then if x is distributed according to $d\lambda_N$, the collection of random variables

$$\{M^{1/2}(p_k(x) - \mathbb{E}[p_k(x)])\}_{k \in \mathbb{N}}$$

converges in the sense of moments to the Gaussian vector with zero mean and covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}\left(M^{1/2}p_k(x), M^{1/2}p_l(x)\right) \\ = kl \oint \oint \Xi(u)\Xi(w)\Psi(u)^{k-1}\Psi(w)^{l-1}\Lambda(u, w) \frac{du}{2\pi\mathbf{i}} \frac{dw}{2\pi\mathbf{i}} + kl \oint \Xi(u)\Psi(u)^{k+l-2}\Psi'(u) \frac{du}{2\pi\mathbf{i}}, \end{aligned}$$

where the u and w -contours enclose V_χ and lies within U , and the u -contour is contained inside the w -contour.

1.4. Asymptotics of multivariate Bessel generating functions for matrix products. To apply Theorems 2.6 and 2.15 to products of random matrices, we choose $\chi_N = \rho = (N-1, \dots, 0)$ as the N -tuples for our multivariate Bessel generating functions. For a positive definite matrix X , denote by $\phi_X(s)$ the multivariate Bessel generating function of its log-spectral measure for the N -tuple ρ . Here, it is the log-spectral measure instead of the spectral measure itself which will behave nicely with respect to multivariate Bessel generating functions; moreover, considering the log-spectral measure is necessary in our later study of Lyapunov exponents.

We then use the key property shown in Lemma 4.2 that if $X_1 = Y_1^* Y_1$ and $X_2 = Y_2^* Y_2$ are unitarily invariant random matrices, then the multivariate Bessel generating function for $X_3 = (Y_1 Y_2)^* (Y_1 Y_2)$ is given by

$$\phi_{X_3}(s) = \phi_{X_1}(s) \phi_{X_2}(s).$$

We conclude that the multivariate Bessel generating function of the log-spectral measure for $X_N := (Y_N^1 \dots Y_N^M)^* (Y_N^1 \dots Y_N^M)$ is given by

$$\phi_N(s) = \psi_X(s)^M,$$

where $\psi_X(s)$ is the generating function of the log-spectral measure of $X_N^1 := (Y_N^1)^* Y_N^1$.

Our second key step is an asymptotic analysis in Theorems 3.9 and 3.4 of the ratio $\frac{\mathcal{B}(s, \lambda)}{\mathcal{B}(\rho, \lambda)}$ near $s = \rho$, which allow us to check in Theorem 4.5 that the measures we study are CLT-appropriate for ρ . This analysis hinges on the following new integral representation of this ratio at $s = \mu_{a,b} := (a, N-1, \dots, \hat{b}, \dots, 0)$ motivated by the main integral formula in [GP15], where $\hat{b}$ means that b is omitted from the expression.

Theorem 3.2. For any $a \in \mathbb{C}$ and $b \in \{N-1, \dots, 0\}$, we have that

$$\frac{\mathcal{B}(\mu_{a,b}, \lambda)}{\mathcal{B}(\rho, \lambda)} = (-1)^{N-b+1} \frac{\Gamma(N-b)\Gamma(b+1)\Gamma(a-N+1)(a-b)}{\Gamma(a+1)} \oint_{\{e^{\lambda_i}\}} \frac{dz}{2\pi\mathbf{i}} \oint_{\{0,z\}} \frac{dw}{2\pi\mathbf{i}} \cdot \frac{z^a w^{-b-1}}{z-w} \cdot \prod_{i=1}^N \frac{w - e^{\lambda_i}}{z - e^{\lambda_i}},$$

where the z -contour encloses the poles at e^{λ_i} and avoids the negative real axis and the w -contour encloses both 0 and the z -contour.

In the previous works [GP15, Cue18a, Cue18b, Cue18c] of Gorin-Panova and Cuenca, representations of normalized symmetric polynomials as single contour integrals were used for the asymptotic analysis. We were not able to find such a representation for the normalized version of the multivariate Bessel functions that we need in this text. Nevertheless, it turns out that we can increase the dimension of the integration cycle and instead use a double contour integral, which remains well-suited for asymptotic analysis. We also remark that our double contour integral in Theorem 3.2 has certain similarities with Okounkov's representation of the correlation kernel for the Schur measure in [Oko01, Theorem 2].

1.5. Multivariate Bessel generating functions and the Cholesky decomposition. When applied to the log-spectral measure of a positive definite Hermitian unitarily invariant random matrix X , the multivariate Bessel generating function for ρ has an interpretation in terms of the Cholesky decomposition of X . Recall that the Cholesky decomposition of a positive definite Hermitian matrix X is a factorization $X = R^* R$ of X into the product of an upper triangular matrix R and its conjugate transpose. In Proposition B.3 we

relate the multivariate Bessel generating function $\phi_X(s)$ for the log-spectral measure of X to the Mellin transform of the diagonal elements of R via

$$(1.1) \quad \phi_X(s) = \mathbb{E} \left[\prod_{k=1}^N R_{kk}^{2(s_k - \rho_k)} \right].$$

This is a translation of [KK16b, Lemma 5.3] of Kieburg-Kosters to our language, and we view it as an analogue of the fact that the distribution of any unitarily-invariant random matrix X is determined by the joint distribution of the diagonal entries $(X_{kk})_{1 \leq k \leq N}$ which was recently applied by Matsumoto-Novak in [MN18] to the study of sums of random matrices, among other things, and which is implicit in the works [OV96, Gor14] of Olshanski-Vershik and Gorin.

We may apply this connection to give an intuitive explanation for the multiplicative property of multivariate Bessel generating functions over products of matrices. If $X_1 = Y_1^* Y_1$ and $X_2 = Y_2^* Y_2$ are independent unitarily invariant random matrices with Cholesky decompositions $X_1 = (R^1)^* R^1$ and $X_2 = (R^2)^* R^2$, then for $X_3 = (Y_1 Y_2)^* Y_1 Y_2$ with Cholesky decomposition $X_3 = (R^3)^* R^3$, we give in Proposition B.2 a direct geometric argument that

$$R_{kk}^3 \stackrel{d}{=} R_{kk}^1 \cdot R_{kk}^2.$$

Combined with (1.1), this gives another proof that $\phi_{X_3}(s) = \phi_{X_1}(s) \phi_{X_2}(s)$.

We also apply this result to give a geometric interpretation of Voiculescu's S -transform. Let $d\lambda$ be a compactly supported measure on $(0, \infty)$, and let λ^N be N -tuples whose empirical measures converge weakly to $d\lambda$. Then we have the following expression for the S -transform of $d\lambda$ in terms of the limiting Cholesky decompositions of unitarily invariant random matrices with fixed spectra λ^N .

Corollary B.4. Let X_N be the $N \times N$ unitarily invariant Hermitian random matrix with spectrum λ^N . For $t \in [0, 1]$, the log- S -transform of the measure $d\lambda$ is given by

$$-\log S_{d\lambda}(t-1) = \lim_{N \rightarrow \infty} \mathbb{E} [2 \log R_{[tN], [tN]}],$$

where $X_N = R^* R$ is the Cholesky decomposition of X_N .

1.6. Relation to the literature. Our results relate to several different lines of work in the literature. In many cases, we are able to prove more general results than those previously known in the same settings, though some of the papers we mention are focused on different applications. We now summarize these connections, beginning with relations to previous results in the random matrix theory literature.

- Global fluctuations for sums and products of finitely many unitarily-invariant random matrices with fixed spectrum were previously studied in some cases using techniques from the Stieltjes transform, free probability, and Schwinger-Dyson equations. For sums, a central limit for global fluctuations was proven using Stieltjes transform techniques in the works [PV07, PS11] of Pastur-Vasilchuk and Pastur-Shcherbina and follows from the theory of second-order freeness in free probability (see the papers [MS06, MSS07, CMSS07] of Mingo-Speicher, Mingo-Śniady-Speicher, and Collins-Mingo-Śniady-Speicher) as described in [MS17, Chapter 5]. A limiting case of our methods recovers the case of sums of random matrices; this approach was followed by Bufetov-Gorin in [BG18a, Section 9.4] and shown to recover the results of [PS11].

For products, a central limit theorem was shown using Stieltjes transform techniques for the product of two unitary matrices with fixed spectrum by Vasilchuk in [Vas16]. For products of positive definite matrices, second-order freeness yields a CLT without an expression for the covariance via the results of Collins-Mingo-Śniady-Speicher in [CMSS07, Theorem 7.9 and Theorem 8.2] and with non-explicit combinatorial formula for the covariance via the results of Arizmendi-Mingo in [AM18]. In [GN15], Guionnet-Novak show that all polynomial linear statistics are asymptotically Gaussian using the Schwinger-Dyson equations, though no explicit formula for the covariance is given and the same techniques do not extend to the log-spectrum. Finally, the recent paper [CO18] of Coston-O'Rourke gives a CLT for linear statistics of eigenvalues of products of Wigner matrices. From this perspective, our Theorem 4.8 yields the first explicit formula for the covariance of the CLT for products of positive definite matrices. Simultaneously, this is the first result showing that this covariance grows logarithmically on short scales.

- Global fluctuations for some classes of matrix models amenable to our techniques have been studied extensively in the literature. For Wigner, Wishart, and Jacobi ensembles, it was shown in the general β setting in the works [DP12, Jia13, Bor14a, Bor14b, BG15a, DP18] of Dumitriu-Paquette, Jiang, Borodin, and Borodin-Gorin that the multilevel eigenvalue structure of these ensembles has Gaussian free field fluctuations. For separable covariance matrices, namely matrices of the form AXB with A, B deterministic and X rectangular Ginibre, a CLT for global fluctuations was shown by Bai-Li-Pan in [BLP16] via Stieltjes transform techniques. Our Theorems 4.12 and 4.13 recover the single level versions of these results. Finally, for a product of finitely many complex Ginibre matrices, the paper [KZ14] of Kuijlaars-Zhang interprets the process of singular values as a biorthogonal ensemble with explicit recurrence relation, which satisfies a global CLT by the results of Breuer-Duits in [BD17].
- In the setting of the M -fold product of $N \times N$ random matrices, the law of large numbers for Lyapunov exponents as $M \rightarrow \infty$ and then $N \rightarrow \infty$ was originally studied by Newman and Isopi-Newman in [New86a, New86b, IN92]. It was then obtained in a free probability context by Kargin and Tucci in [Kar08, Tuc10]. Our techniques recover these law of large numbers results.

In the more recent works [For13, ABK14] of Forrester and Akemann-Burda-Kieburg, a LLN and CLT were shown for square Ginibre ensembles with fixed N and $M \rightarrow \infty$, and Forrester studied in [For15] their asymptotics as $N \rightarrow \infty$. In [Red16], Reddy showed a general LLN and CLT for Lyapunov exponents of unitarily invariant ensembles at fixed N . Finally, in [KK16b], Kieburg-Kosters gave a different proof for a subset of these ensembles which they call polynomial ensembles of derivative type (which corresponds to cases where the multivariate Bessel generating function is multiplicative). However, none of these papers studies the CLT in our regime, where $M, N \rightarrow \infty$ simultaneously.

- Local statistics of Lyapunov exponents for products of square Ginibre matrices are studied for $M, N \rightarrow \infty$ simultaneously by Akemann-Burda-Kieburg and Liu-Wang-Wang in [ABK18] and [LWW18]. The local limits are shown to have picket fence statistics when $M/N \rightarrow \infty$, sine kernel statistics when $M/N \rightarrow 0$, and an interpolation between the two in the intermediate regime when $M \sim N$. In particular, their results imply that the limits $N \rightarrow \infty$ and $M \rightarrow \infty$ do not commute on the local scale for Ginibre matrices, as earlier predicted heuristically by the same authors in [ABK14, Section 5].

Their methods rely an expression for the process of squared singular values as a determinantal point process with explicit correlation kernel, and we believe this expression may be modified to obtain similar results for rectangular Ginibre matrices. When compared to our Theorem 4.17, this yields a transition between the sine kernel on local scales and a Gaussian process with white noise on the global scale.

- In the study of dynamical isometry for deep neural networks, singular values of products of many i.i.d. copies of large random matrices appear as Jacobians of newly initialized networks in the work of many researchers in the papers [SMG14, PSG17, PSG18, CPS18, XBSD⁺18, TWJ⁺18, LQ18], which study different network architectures which lead to singular values closely clustered around 1. We mention also the recent paper [HN18] of Hanin-Nica, which proves Gaussianity for the first exponential moment of the Lyapunov exponents for a specific family of matrix products originating in the study of deep neural networks. While these works consider orthogonally-invariant or general real matrix ensembles (as opposed to unitarily-invariant complex matrix ensembles), it is believed that the structure of fluctuations is similar in both cases, allowing us to make predictions for the behavior of fluctuations in this setting.

Our technique via multivariate Bessel generating functions also has many relations to the literature, which we now discuss.

- The idea that products of unitarily-invariant random matrices are related to products of the corresponding multivariate Bessel functions dates back to at least Macdonald and Zinn-Justin in [Mac95, Section VII.4] and [ZJ99], where it was used to give an heuristic proof of the multiplicativity of Voiculescu's S -transform. Our work may be viewed as a rigorous realization and generalization of this approach. It rests on the analytic continuation of the key functional equation for characters of $\mathfrak{gl}_N$ used in Lemma 4.2, which holds for any compact group and whose analytic continuation was studied by Berezin-Gelfand in [BG56] and Helgason in [Hel84]. We mention also the recent

work [BGS18] of Borodin-Gorin-Strahov, where the eigenvalues of products of Jacobi ensembles were shown to form a continuous limit of a Schur process with certain specializations.

When applied to log-spectral measures of unitarily invariant random matrix ensembles, our multivariate Bessel generating functions correspond after a change of variables to the spherical transform as surveyed in [Hel84]. The spherical transform is used by Kieburg-Kosters in [KK16a, KK16b] to derive correlation kernels for products of random matrices drawn from so-called polynomial ensembles and to obtain a CLT for Lyapunov exponents of fixed size matrices when the polynomial ensembles are of derivative type. In our language, this is an analogue of Theorem 2.15 in the special case where N is fixed, $M \rightarrow \infty$, and the multivariate Bessel generating function is multiplicative and hence $\Lambda(u, w) = 0$. In the general β setting, the paper [GM17] of Gorin-Marcus uses the fact that such spherical transforms are scaling limits of discrete versions originating from the theory of Macdonald symmetric functions.

- Our main technical tools Theorems 2.6 and 2.15 are in the spirit of the previous works [BG15b, BG18a, BG18b, Hua18] of Bufetov-Gorin and Huang which together established similar results on law of large numbers and central limit theorems for measures whose Schur/Jack generating functions admit nice expansions near the point $(1, 1, \dots, 1)$. Our theorems may be interpreted as a partial generalization of the (continuous limit of) the Schur case of these results to expansions near arbitrary N -tuples χ , although this introduces significant complications in the proofs. In particular, the log-derivatives of the multivariate Bessel generating function evaluated at this point may depend on the variables in which the derivatives were taken, which was not true for evaluation of Schur functions at $(1, 1, \dots, 1)$. Accounting for this in the computation of moments and covariance in Theorems 2.6 and 2.15 leads to the integrals along contours around the support V_χ of $d\chi$ in their statements.

We mention also the series of papers [BL17, Li18a, Li18b] by Boutillier-Li and Li which give analogues of these results for expansions of Schur generating functions near a point $(x_1, \dots, x_n, x_1, \dots, x_n, \dots, x_1, \dots, x_n)$ with periodic coordinates, though we note that those papers obtain asymptotics of Schur generating functions using different techniques.

- The integral formula (3.2) in Theorem 3.2 for evaluations of multivariate Bessel functions near ρ and its Schur analogue (3.11) were inspired by the alternate proof of [GP15, Theorem 3.6] discussed in the remark immediately after its statement. The Schur version of this formula is also used by Cuenca-Gorin in [CG18, Theorem 3.4] to study the q -analogue of the Gelfand-Tsetlin branching graph.

1.7. Organization of the paper. The remainder of this paper is organized as follows. In Section 2, we introduce the multivariate Bessel generating functions and prove our main technical results, Theorems 2.6 and 2.15, which give central limit theorems for global fluctuations of measures with nice multivariate Bessel generating functions. In Section 3, we give a new contour integral formula for certain multivariate Bessel functions in Theorem 3.2 and use it to analyze their asymptotics in Theorems 3.9 and 3.4. In Section 4, we apply the results of the previous two sections to prove our main results, Theorems 4.8, 4.14, 4.16, and 4.17 on the convergence of height functions to Gaussian random fields. Finally, in Section 5, we present an extension of our technique to obtain a 2-D Gaussian field from products of different numbers of random matrices in Theorem 5.1. For the convenience of the reader, all definitions and notations will be given again in the following sections.

Acknowledgments. We would like to thank Andrew Ahn, Alexei Borodin, Maurice Duits, Octavio Arizmendi Echegaray, Alice Guionnet, Victor Kleptsyn, James Mingo, Roland Speicher, and Alex Zhai for useful discussions. Both authors were supported by the NSF grants DMS-1664619 and DMS-1664650. V. G. was partially supported by the NEC Corporation Fund for Research in Computers and Communications and by the Sloan Research Fellowship. Y. S. was supported by a Junior Fellow award from the Simons Foundation and NSF Grant DMS-1701654/2039183. The authors also thank the organizers of the Park City Mathematics Institute research program on Random Matrix Theory and the MATRIX workshop “Non-Equilibrium Systems and Special Functions,” where part of this work was completed.

2. CENTRAL LIMIT THEOREM VIA MULTIVARIATE BESSEL GENERATING FUNCTIONS

2.1. Multivariate Bessel generating functions. For a set of variables $a = (a_1, \dots, a_N)$, we denote the Vandermonde determinant by $\Delta(a) := \prod_{1 \leq i < j \leq N} (a_i - a_j)$. For sets of complex variables $a = (a_1, \dots, a_N)$

and $b = (b_1, \dots, b_N)$ with all $\{a_i\}$ and $\{b_i\}$ distinct, the multivariate Bessel function is given by

$$(2.1) \quad \mathcal{B}(a, b) := \Delta(\rho) \frac{\det(e^{a_i b_j})_{i,j=1}^N}{\Delta(a)\Delta(b)},$$

where $\rho = (N-1, \dots, 0)$, and it admits analytic continuation to all values of a and b . For a probability measure $d\mu_N$ on $W^N := \{(x_1, \dots, x_N) \in \mathbb{R}^N \mid x_1 \geq \dots \geq x_N\}$ and an N -tuple $\chi \in W^N$, define its multivariate Bessel generating function $\phi_{\chi, N}(s) := \phi_{\chi, N}(s_1, \dots, s_N)$ by

$$(2.2) \quad \phi_{\chi, N}(s) := \int \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi, x)} d\mu_N(x).$$

Definition 2.1. We say that a measure $d\mu_N$ is χ -smooth if the defining integral of $\phi_{\chi, N}(s)$ converges absolutely and uniformly on a complex neighborhood of $[\chi_N, \chi_1]^N$.

Remark. In our applications in later sections, we will take $\chi = \rho$. For convenience, in those applications we will omit the χ 's in the notation and write simply $\phi_N(s)$ and smooth instead of $\phi_{\rho, N}(s)$ and ρ -smooth.

Define the operators

$$(2.3) \quad D_k := \Delta(s)^{-1} \circ \sum_{i=1}^N \partial_{s_i}^k \circ \Delta(s)$$

and the moments $p_k(x) := x_1^k + \dots + x_N^k$. We show now that repeated applications of the operators D_k yield mixed moments of a smooth measure.

Lemma 2.2. For any $k_1, \dots, k_l \geq 1$, if $d\mu_N$ is χ -smooth and x is $d\mu_N$ -distributed, then $\mathbb{E}[p_{k_1}(x) \cdots p_{k_l}(x)]$ is finite and equals

$$\mathbb{E}[p_{k_1}(x) \cdots p_{k_l}(x)] = D_{k_1} \cdots D_{k_l} \phi_{\chi, N}(\chi).$$

Proof. Choose $\delta > 0$ small so that for s in a neighborhood of $[\chi_N, \chi_1]^N$, the defining integral of $\phi_{\chi, N}$ is convergent at $s + \delta\varepsilon$ for all 2^N choices of vector ε with entries in $\{\pm 1\}$. We therefore see that on this neighborhood we have

$$\begin{aligned} \sum_{\varepsilon_i \in \{\pm 1\}} \Delta(s + \delta\varepsilon) \phi_{\chi, N}(s + \delta\varepsilon) &= \int \frac{\Delta(\chi)}{\det(e^{\chi_i x_j})_{i,j=1}^N} \sum_{\varepsilon_i \in \{\pm 1\}} \sum_{\sigma \in S_N} (-1)^\sigma \prod_{i=1}^N e^{s_i x_{\sigma(i)} + \delta\varepsilon_i x_{\sigma(i)}} d\mu_N(x) \\ &= \int \frac{\Delta(\chi)}{\det(e^{\chi_i x_j})_{i,j=1}^N} \prod_{i=1}^N [2 \cosh(\delta x_i)] \sum_{\sigma \in S_N} (-1)^\sigma \prod_{i=1}^N e^{s_i x_{\sigma(i)}} d\mu_N(x) \\ &= \Delta(s) \int \prod_{i=1}^N [2 \cosh(\delta x_i)] \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi, x)} d\mu_N(x). \end{aligned}$$

Notice now that

$$\sum_{\varepsilon_i \in \{\pm 1\}} \Delta(s + \delta\varepsilon) \phi_{\chi, N}(s + \delta\varepsilon)$$

vanishes at $s_i = s_j$, meaning that

$$(2.4) \quad \sum_{\varepsilon_i \in \{\pm 1\}} \frac{\Delta(s + \delta\varepsilon)}{\Delta(s)} \phi_{\chi, N}(s + \delta\varepsilon) = \int \prod_{i=1}^N [2 \cosh(\delta x_i)] \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi, x)} d\mu_N(x)$$

admits an analytic continuation to and converges absolutely on a neighborhood of $[\chi_N, \chi_1]^N$ for some $\delta > 0$. Notice now that $|p_{k_1}(x) \cdots p_{k_l}(x)| < \prod_{i=1}^N 2 \cosh(\delta x_i)$ outside of a compact set, which implies that

$$\left| p_{k_1}(x) \cdots p_{k_l}(x) \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi, x)} \right| < \left| \prod_{i=1}^N [2 \cosh(\delta x_i)] \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi, x)} \right|$$

outside of a compact set. Therefore, the absolute convergence of (2.4) together with the eigenrelation

$$D_k \mathcal{B}(s, x) = p_k(x) \mathcal{B}(s, x)$$

implies by dominated convergence that for s in some neighborhood of $[\chi_N, \chi_1]^N$, we have

$$D_{k_1} \cdots D_{k_l} \phi_{\chi, N}(s) = \int p_{k_1}(x) \cdots p_{k_l}(x) \frac{\mathcal{B}(s, x)}{\mathcal{B}(\chi, x)} d\mu_N(x),$$

Plugging in $s = \chi$ then yields

$$\mathbb{E}[p_{k_1}(x) \cdots p_{k_l}(x)] = \int p_{k_1}(x) \cdots p_{k_l}(x) d\mu_N = D_{k_1} \cdots D_{k_l} \phi_{\chi, N}(\chi). \quad \square$$

2.2. Statement of the results. Let $\chi_N \in W^N$ be a sequence of N -tuples such that we have the weak convergence of measures

$$\frac{1}{N} \sum_{i=1}^N \delta_{\chi_{N,i}/N} \rightarrow d\chi$$

for some compactly supported measure $d\chi$ on $\mathbb{R}$. Let $\phi_{\chi, N}(s) := \phi_{\chi, N}(s_1, \dots, s_N)$ be the multivariate Bessel generating function of a sequence of smooth measures $d\mu_N$. We now define conditions on $\phi_{\chi, N}(s)$ which imply a limit shape and Gaussian fluctuations for these measures. For a subset $I \subset \{1, \dots, N\}$, define $s^I := (s_i)_{i \in I}$ and

$$\phi_{\chi, N}^I(s) := \phi_{\chi, N}(s) |_{s_j = \chi_j, j \notin I},$$

where we note that $\phi_{\chi, N}^I(s)$ is a function of s^I alone. For a function $f(s)$ on $\mathbb{C}^N$ and some $r \in \mathbb{C}^N$, we will adopt the notation that $\partial_{r_i} f(r) = \left. \frac{\partial}{\partial s_i} f(s) \right|_{s=r}$.

Definition 2.3. We say that the measures $d\mu_N$ are LLN-appropriate for χ_N if they are χ_N -smooth and there exist a compact set $V_\chi \subset \mathbb{R}$ and a holomorphic function Ψ on an open complex neighborhood U of V_χ such that $\chi_{N,i}/N \in V_\chi$ for all N and for each fixed integer $k > 0$, uniformly in I with $|I| = k$, $i \in I$, and $r^I \in U^{|I|}$, we have

$$\lim_{N \rightarrow \infty} \left| \frac{1}{N} \partial_{r_i} [\log \phi_{\chi, N}^I(rN)] - \Psi(r_i) \right| = 0.$$

Definition 2.4. We say that measures $d\mu_N$ are CLT-appropriate for χ_N if they are LLN-appropriate for χ_N with respect to some function Ψ and sets $V_\chi \subset \mathbb{R}$ with neighborhood U and there exists a holomorphic function Λ on U^2 such that for each fixed integer $k > 0$, uniformly in I with $|I| = k$, $r^I \in U^{|I|}$, and distinct $i, j \in I$, we have

$$\lim_{N \rightarrow \infty} |\partial_{r_i} \partial_{r_j} [\log \phi_{\chi, N}^I(rN)] - \Lambda(r_i, r_j)| = 0.$$

Define the Cauchy transform of the measure $d\chi$ by

$$(2.5) \quad \Xi(u) := \int \frac{1}{u - x} d\chi(x)$$

as a holomorphic function on $\mathbb{C} \setminus V_\chi$.

Theorem 2.5 (Law of large numbers). If the measures $d\mu_N$ are LLN-appropriate for χ_N , then we have the following convergence in probability

$$(2.6) \quad \lim_{N \rightarrow \infty} \frac{1}{N} p_k(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(x)] = \mathfrak{p}_k := \frac{1}{k+1} \oint \left(\Xi(u) + \Psi(u) \right)^{k+1} \frac{du}{2\pi i},$$

where the u -contour encloses V_χ and lies within U . In addition, the random measures $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ with x distributed according to $d\mu_N$ converge in probability to a deterministic compactly supported measure $d\mu$ with $\int x^k d\mu(x) = \mathfrak{p}_k$.

Theorem 2.6 (Central limit theorem). If the measures $d\mu_N$ are CLT-appropriate for χ_N , then the collection of random variables

$$\{p_k(x) - \mathbb{E}[p_k(x)]\}_{k \in \mathbb{N}}$$

converges in the sense of moments to the Gaussian vector with zero mean and covariance

$$(2.7) \quad \lim_{N \rightarrow \infty} \text{Cov}(p_k(x), p_l(x)) = \mathfrak{Cov}_{k,l} := \oint \oint \left(\Xi(u) + \Psi(u) \right)^l \left(\Xi(w) + \Psi(w) \right)^k \left(\frac{1}{(u-w)^2} + \Lambda(w, u) \right) \frac{dw}{2\pi i} \frac{du}{2\pi i},$$

where the u and w -contours enclose V_χ and lie within U , and the u -contour is contained inside the w -contour.

Remark. In Theorems 2.5 and 2.6, the coordinates of χ_N are allowed to repeat. In particular, we may choose $\chi_N = (0, \dots, 0)$, in which case they are a continuous version of the results of [BG15b, BG18a] and are closely related to the results of [MN18].

In our applications, we will use the following computation of $\Xi(u)$ for $d\chi(x) = \mathbf{1}_{[0,1]}dx$.

Lemma 2.7. If $d\chi(x) = \mathbf{1}_{[0,1]}dx$, we have that $\Xi(u) = \log(u/(u-1))$ on $\mathbb{C} \setminus [0, 1]$.

Proof. By direct computation of $\int_0^1 \frac{1}{u-x} dx$. □

2.3. The symmetrization procedure on analytic functions. In what follows, we will make extensive use of the following properties of symmetrizations of analytic functions.

Definition 2.8. Let $f(s^I)$ be a function in variables $(s_i)_{i \in I}$. For $K = \{k_1, \dots, k_l\} \subset I$, we define its symmetrization over $(s_k)_{k \in K}$ by

$$\text{Sym}_K f(s^I) := \frac{1}{l!} \sum_{\sigma \in S_l} f((s_i^\sigma)_{i \in I}),$$

where for $\sigma \in S_l$, we define

$$s_i^\sigma = \begin{cases} s_i & i \notin K \\ s_{k_{\sigma(m)}} & i = k_m \end{cases}.$$

Lemma 2.9. If $f(s^I)$ is an analytic function of s^I on $U^{|I|}$ for a complex domain $U \subset \mathbb{C}$, for any $A \subset I$ and $i \in I \setminus A$, the symmetrization over $K := \{i\} \cup A$ given by

$$\text{Sym}_K \left[\prod_{a \in A} \frac{1}{s_i - s_a} f(s^I) \right]$$

is an analytic function of s^I on $U^{|I|}$. Further, if $f_t(s^I)$ for $t = 1, 2, \dots$ are a sequence of analytic functions uniformly converging to 0 on $U^{|I|}$, then the sequence of functions

$$\text{Sym}_K \left[\prod_{a \in A} \frac{1}{s_i - s_a} f_t(s^I) \right]$$

converges to 0 uniformly on compact subsets of $U^{|I|}$.

Proof. Denote by $g(s^I)$ the symmetrized function, which is meromorphic in s^I on $U^{|I|}$. Notice that the poles of $g(s^I)$ are contained within the hyperplanes $\{s^I \in U^{|I|} \mid s_{k_1} = s_{k_2}\}$ for $k_1, k_2 \in K$. The complement of the set

$$V := \{s^I \in U^{|I|} \mid s_{k_1} = s_{k_2}, s_k \text{ distinct and not equal to } s_{k_1} \text{ for } k \notin \{k_1, k_2\}\}$$

in the union of these hyperplanes has codimension 2. Because the set of poles is either empty or has codimension 1 by Riemann's second extension theorem (see [GR84, Theorem 7.1.2]), it suffices for us to check that the poles of $g(s^I)$ avoid V . For any $r^I \in V$ with $r_{k_1} = r_{k_2} = z$, we see that $g(s^I)$ is evidently holomorphic near r^I in the variables $(s_i)_{i \in I \setminus \{k_1, k_2\}}$. Notice now that we may rewrite $g(s^I)$ as

$$(2.8) \quad g(s^I) = \text{Sym}_K \left[\frac{1}{s_{k_1} - s_{k_2}} \frac{1}{s_{k_1} - s_i} \prod_{k \in K \setminus \{i, k_1, k_2\}} \frac{1}{s_{k_1} - s_k} f(\tau(s^I)) \right] \\ = \frac{1}{2} \text{Sym}_K \left[\frac{1}{s_{k_1} - s_{k_2}} \left(\prod_{k \in K \setminus \{k_1, k_2\}} \frac{1}{s_{k_1} - s_k} f(\tau(s^I)) - \prod_{k \in K \setminus \{k_1, k_2\}} \frac{1}{s_{k_2} - s_k} f(\tau'(s^I)) \right) \right],$$

where τ exchanges s_i with s_{k_1} and τ' exchanges s_i with s_{k_2} . Notice that

$$\prod_{k \in K \setminus \{k_1, k_2\}} \frac{1}{s_{k_1} - s_k} f(\tau(s^I)) \quad \text{and} \quad \prod_{k \in K \setminus \{k_1, k_2\}} \frac{1}{s_{k_2} - s_k} f(\tau'(s^I))$$

are both analytic near $s^I = r^I$ and are given by exchanging s_{k_1} with s_{k_2} , which implies that the argument of the symmetrization in (2.8) and hence $g(s^I)$ are analytic in s_{k_1} and s_{k_2} near r^I . Thus $g(s^I)$ is analytic in all variables near r^I , completing the proof of the first claim.

For the second claim, for any compact subset $V' \subset U^{|I|}$, choose $|K|$ nested contours $(\gamma_k)_{k \in K}$ contained in U and containing the coordinate projections of V' . Notice that for any partition $K = \{i\} \sqcup A$, the function $\prod_{a \in A} \frac{1}{u_i - u_a}$ is uniformly bounded for $(u_k)_{k \in K}$ with $u_k \in \gamma_k$. Denoting by $g_t(s^I)$ the symmetrized function corresponding to $f_t(s^I)$, this implies that $g_t \rightarrow 0$ uniformly on $\{s \in U \mid s_k \in \gamma_k\}$. Applying the Cauchy integral formula $|K|$ times with the $|K|$ nested contours $(\gamma_k)_{k \in K}$ then implies that g_t converges to 0 uniformly on V' , as desired. $\square$

2.4. A graphical calculus. The purpose of this section is to produce an expansion for $m \leq n$ of

$$\frac{D_{k_m} \cdots D_{k_n} \phi_{\chi, N}(s)}{\phi_{\chi, N}(s)}$$

into the sum of terms. Each term will be associated with a *forest* tagged with some combinatorial data, where many terms can have the same index. The proofs of Theorems 2.5 and Theorem 2.6 will identify terms associated to specific forests as leading order, allowing us to only compute the values of those terms. In what follows, we specify this expansion precisely via an inductive construction in m .

In what follows, for each $m \in \{1, \dots, n\}$, we will define certain *type m terms* appearing in our expansion by downwards induction on m . Such type m terms will be associated with a rooted forest F on the vertex set $\{m, \dots, n\}$ with either single or bold edges along with an assignment to each tree T in F of a set of indices $K_T \subseteq \{1, \dots, N\}$, a LLN weight $W_{T, L} \geq 0$ and a CLT weight $W_{T, C} \geq 0$. Let $\mathcal{F}_{m, n}$ be the set of such forests with associated data. This association will have the property that if $F = T_1 \sqcup \cdots \sqcup T_l \in \mathcal{F}_{m, n}$ is a division of the forest into trees, then the term takes the form

$$X_1 \cdots X_l,$$

where each X_i is associated with T_i . We now proceed to define the notion of a type m term.

Call sets of indices $A_1, \dots, A_n \subseteq \{1, \dots, N\}$, indices $j_1, \dots, j_n \in \{1, \dots, N\}$, and (possibly-empty) vectors of non-negative integers $B_1, \dots, B_n$ and $C = (c_1, \dots, c_{n-1})$ a *good label* if $j_i \notin A_i$, $B_i = (b_i^1, \dots, b_i^{i-1})$ with $b_j^i > 0$, $c_n = 0$, and $k_i = |A_i| + \sum_j b_j^i + c_i$. We define a type n term to be an expression of the form

$$\text{Sym}_{\{j_n\} \cup A_n} \left[\prod_{a \in A_n} \frac{1}{s_{j_n} - s_a} \right] [\partial_{j_n}^{b_n^1} \log \phi_{\chi, N}(s)] \cdots [\partial_{j_n}^{b_n^{i-1}} \log \phi_{\chi, N}(s)]$$

where A_n, j_n, B_n come from a good label. We associate to this type n term the forest in $\mathcal{F}_{n, n}$ consisting of the single vertex tree T rooted at n with index set $K_T := \{j_n\} \cup A_n$ and LLN and CLT weights $W_L = W_C = \sum_i (b_i^n - 1) + |A_n|$. It consists of a single multiplicand associated to T .

For $m < n$, we define type m terms by downwards induction on m . Let $X_1 \cdots X_l$ be a type $m+1$ term corresponding to some good label and associated with a forest $F \in \mathcal{F}_{m+1, n}$ consisting of trees $T_1, \dots, T_l$ associated to $X_1, \dots, X_l$, respectively. For some $c_m^1, \dots, c_m^l \geq 0$ such that $c_m^1 + \cdots + c_m^l = c_m$ and where A_i, j_i, B_i, c_i come from the same good label as the type $m+1$ term, define

$$(2.9) \quad K'_T := \{j_m\} \cup A_m \cup \bigcup_{i: c_m^i > 0} K_{T_i}.$$

We define a type m term to be an expression of the form

$$(2.10) \quad \left[\prod_{i: c_m^i = 0} X_i \right] \cdot \text{Sym}_{K'_T} \left[\left[\prod_{a \in A_m} \frac{1}{s_{j_m} - s_a} \right] \left[\prod_{i: c_m^i > 0} \partial_{j_m}^{b_m^i} [X_i] \right] [\partial_{j_m}^{b_m^1} \log \phi_{\chi, N}(s)] \cdots [\partial_{j_m}^{b_m^{i-1}} \log \phi_{\chi, N}(s)] \right].$$

We associate this term to the forest $F' \in \mathcal{F}_{m, n}$ consisting of $\{T_i \mid c_m^i = 0\}$ and a new tree T with root at m with edges to the root of each tree in $\{T_i \mid c_m^i > 0\}$. Let the edge between m and the root of T_i with $c_m^i > 0$

be single if $j_m \notin K_{T_i}$ and bold otherwise. Assign to T the index set K'_T and LLN and CLT weight given by

$$W'_L := |A_m| + \sum_i (b_i^m - 1) + c_m + \sum_{\partial_{j_m} \text{ applied to } X_i} W_{T_i, L}$$

$$W'_C := |A_m| + \sum_i (b_i^m - 1) + c_m + \sum_{\partial_{j_m} \text{ applied to } X_i} (W_{T_i, C} + \mathbf{1}_{j_m \notin K_{T_i}}).$$

Finally, associate T_i to X_i and T to the multiplicand

$$\text{Sym}_{K'_T} \left[\left[\prod_{a \in A_m} \frac{1}{s_{j_m} - s_a} \right] \left[\prod_{i: c_i^m > 0} \partial_{j_m}^{c_i^m} [X_i] \right] [\partial_{j_m}^{b_1^m} \log \phi_{\chi, N}(s)] \cdots [\partial_{j_m}^{b_{l_m}^m} \log \phi_{\chi, N}(s)] \right].$$

In Proposition 2.10, we show that these terms are sufficiently rich to expand the action of our differential operators in. Lemma 2.11 and Corollary 2.12 then prove properties of this expansion.

Proposition 2.10. The quantity

$$\frac{D_{k_m} \cdots D_{k_n} \phi_{\chi, N}(s)}{\phi_{\chi, N}(s)}$$

is a linear combination with coefficients independent of N of type- m terms, where the number of terms associated to any $F \in \mathcal{F}_{m, n}$ is independent of N .

Remark. Before giving a proof, for the reader's convenience we illustrate Proposition 2.10 explicitly for $n = 2$. Direct computation shows that

$$(2.11) \quad \frac{D_{k_1} D_{k_2} \phi_{\chi, N}(s)}{\phi_{\chi, N}(s)} = \frac{1}{\phi_{\chi, N}(s)} \sum_{j_1, j_2=1}^N \sum_{\substack{A_1 \subset \{1, \dots, N\} \\ j_1 \notin A_1}} \sum_{\substack{A_2 \subset \{1, \dots, N\} \\ j_2 \notin A_2}} \binom{k_1}{|A_1|} \binom{k_2}{|A_2|} \left[\prod_{a \in A_1} \frac{1}{s_{j_1} - s_a} \right]$$

$$\partial_{s_{j_1}}^{k_1 - |A_1|} \left[\phi_{\chi, N}(s) \frac{1}{\phi_{\chi, N}(s)} \left[\prod_{a \in A_2} \frac{1}{s_{j_2} - s_a} \right] \partial_{s_{j_2}}^{k_2 - |A_2|} \phi_{\chi, N}(s) \right].$$

Proposition 2.10 associates the symmetrized summands of this expression to forests on $\{1, 2\}$. Terms where $\partial_{s_{j_1}}^{k_1 - |A_1|}$ is applied only to the first factor $\phi_{\chi, N}(s)$ on the second line of (2.11) correspond to the forest consisting of two 1-vertex trees $T_1 = \{1\}$ and $T_2 = \{2\}$ with $K_{T_1} = \{j_1\} \cup A_1$ and $K_{T_2} = \{j_2\} \cup A_2$. The remaining terms correspond to the forest consisting of the single tree $T = \{1, 2\}$ rooted at 1, where the edge from 1 to 2 is single if $j_1 \notin \{j_2\} \cup A_2$ and bold if $j_1 \in \{j_2\} \cup A_2$. In this case, $K_T = \{j_1\} \cup \{j_2\} \cup A_1 \cup A_2$.

If $k_1 = k_2 = 1$, and $F \in \mathcal{F}_{1, 2}$ is the forest $\{\{1\}, \{2\}\}$ with given K_{T_1}, K_{T_2} , there are $|K_{T_1}| \cdot |K_{T_2}|$ possibilities for j_1, j_2, A_1, A_2 . Neither the number of terms in the expansion (2.11) associated to this data nor the coefficients of those terms depend on N . However, we remark that K_{T_1} and K_{T_2} can range over subsets of $\{1, \dots, N\}$ of size at most 1, meaning the number of possible choices for F does depend on N .

Proof of Proposition 2.10. We proceed by downwards induction on m . For $m = n$, by the identity of differential operators

$$(2.12) \quad \Delta(s)^{-1} \partial_{s_i}^{k_n} \Delta(s) = \sum_{h=0}^{k_n} \binom{k_n}{h} \sum_{\substack{A_n \subset \{1, \dots, N\} \\ |A_n|=h, i \notin A_n}} \left[\prod_{a \in A_n} \frac{1}{s_i - s_a} \right] \partial_{s_i}^{k_n - h},$$

we conclude that

$$D_{k_n} \phi_{\chi, N}(s) = \Delta(s)^{-1} \sum_{i=1}^N \partial_{s_i}^{k_n} \Delta(s) \phi_{\chi, N}(s)$$

$$= \sum_{i=1}^N \sum_{\substack{A_n \subset \{1, \dots, N\} \\ i \notin A_n}} \binom{k_n}{|A_n|} \left[\prod_{a \in A_n} \frac{1}{s_i - s_a} \right] \partial_{s_i}^{k_n - |A_n|} \phi_{\chi, N}(s).$$

Because $D_{k_n} \phi_{\chi, N}(s)$ is symmetric in s , we may replace the summand of this double summation by its symmetrization over $\{s_j \mid j \in \{i\} \cup A_n\}$ multiplied by a constant independent of N . Consequently, by the Leibnitz rule and the identity

$$\partial_{s_j} \phi_{\chi, N}(s) = [\partial_{s_j} \log \phi_{\chi, N}(s)] \phi_{\chi, N}(s),$$

this symmetrized summand is exactly a linear combination of terms of type n with coefficients independent of N . Since $k_n = |A_n| + \sum_j b_j^n$ by construction, the number of such terms associated with a fixed choice of index set $K_T = \{j_n\} \cup A_n$ and LLN and CLT weights $W_L = W_C$ is a function of k_n and therefore also independent of N .

Suppose now that the claim is true for some $m+1 \leq n$. It suffices to check that for any type $m+1$ term $X_1 \cdots X_l$, the expression

$$\frac{D_{k_m}[X_1 \cdots X_l \phi_{\chi, N}(s)]}{\phi_{\chi, N}(s)}$$

is a linear combination of type m terms. By (2.12), we find that

$$D_{k_m}[X_1 \cdots X_l \phi_{\chi, N}(s)] = \sum_{i=1}^N \sum_{\substack{A_m \subset \{1, \dots, N\} \\ i \notin A_m}} \binom{k_m}{|A_m|} \left[\prod_{a \in A_m} \frac{1}{s_i - s_a} \right] \partial_{s_i}^{k_m - |A_m|} [X_1 \cdots X_l \phi_{\chi, N}(s)].$$

Applying the Leibnitz rule to distribute the derivative and symmetrizing over $\{s_i \mid i \in K'_T\}$ with K'_T from (2.9) the portion of the resulting summand which excludes terms X_i whose s_i derivative are not taken, we obtain a decomposition into a linear combination of type m terms. By the inductive hypothesis, the number of type $m+1$ terms $X_1 \cdots X_l$ associated to $F \in \mathcal{F}_{m+1, n}$ is independent of N and their coefficients are independent of N . In addition, if a type m term is associated to $F' \in \mathcal{F}_{m, n}$, it comes from a type $m+1$ term $F' \in \mathcal{F}_{m+1, n}$ via the procedure just described, and the number of possible choices of F' is independent of N . Since the coefficients and number of type m terms associated to F by the procedure above are independent of N , we conclude that the resulting overall decomposition into type m terms again has coefficients independent of N and number of terms associated to $F \in \mathcal{F}_{m, n}$ independent of N , completing the induction. $\square$

Lemma 2.11. If a term $X_1 \cdots X_l$ corresponds to a forest with trees $\{T_1, \dots, T_l\}$ and index sets $K_{T_1}, \dots, K_{T_l}$, then letting

$$X_t^I(s) := X_t(s)|_{s_j = \chi_j, j \notin I}, \quad t \in \{1, \dots, l\},$$

for

$$I = K_{T_1} \cup \dots \cup K_{T_l},$$

each $X_t^I(s)$ with $t \in \{1, \dots, l\}$ satisfies

- (a) $X_t^I(s)$ is symmetric in $\{s_i\}_{i \in K_{T_t}}$;
- (b) $X_t^I(rN)$ is analytic in r^I ;
- (c) if $d\mu_N$ is LLN-appropriate for χ_N , $N^{W_{T_t, L}} X_t^I(rN)$ converges as $N \rightarrow \infty$ to an analytic function of r^I on $U^{|I|}$;
- (d) if $d\mu_N$ is CLT-appropriate for χ_N , $N^{W_{T_t, C}} X_t^I(rN)$ converges as $N \rightarrow \infty$ to an analytic function of r^I on $U^{|I|}$;
- (e) if $d\mu_N$ is CLT-appropriate for χ_N and $i \in I \setminus K_{T_t}$, $N^{W_{T_t, C}+1} \partial_{r_i} X_t^I(rN)$ converges as $N \rightarrow \infty$ to an analytic function of r^I on $U^{|I|}$;
- (f) the number of appearances of $\log \phi_{\chi, N}^I(s)$ in $X_t^I(s)$ is given by $\sum_{i \in T_t} k_i - W_{T_t, L}$.

Proof. We consider the expansion of

$$\frac{D_{k_m} \cdots D_{k_n} \phi_{\chi, N}(s)}{\phi_{\chi, N}(s)}$$

into terms indexed by forests on $\{m, \dots, n\}$ and apply downward induction on m from n to 1. For $m = n$, we consider terms of the form

$$X_t^I(s) = \text{Sym}_{\{j_n\} \cup A_n} \left[\prod_{a \in A_n} \frac{1}{s_{j_n} - s_a} \right] [\partial_{j_n}^{b_1^n} \log \phi_{\chi, N}^I(s)] \cdots [\partial_{j_n}^{b_{l_n}^n} \log \phi_{\chi, N}^I(s)].$$

Claims (a) and (b) follow from the definition and an application of Lemma 2.9. For (c) and (d), we see that

$$\begin{aligned} N^{W_{T_t}, C} X_t^I(rN) &= N^{W_{T_t}, L} X_t^I(rN) \\ &= \text{Sym}_{\{j_n\} \cup A_n} \left[\prod_{a \in A_n} \frac{1}{r_{j_n} - r_a} \right] [N^{-1} \partial_{r_{j_n}}^{b_{j_n}^n} \log \phi_{\chi, N}^I(rN)] \cdots [N^{-1} \partial_{r_{j_n}}^{b_{j_n}^n} \log \phi_{\chi, N}^I(rN)]. \end{aligned}$$

By Definition 2.3, we see that $\frac{1}{N} \partial_{r_i} \log \phi_{\chi, N}^I(rN)$ converges to an analytic function of r^I , so (c) and (d) both follow from Lemma 2.9. For (e), because $i \notin K_{T_t}$, we see that ∂_{r_i} and $\text{Sym}_{K_{T_t}}$ commute, meaning that

$$N^{W_{T_t}, C+1} \partial_{r_i} X_t^I(rN) = \text{Sym}_{\{j_n\} \cup A_n} \left[\prod_{a \in A_n} \frac{1}{r_{j_n} - r_a} \right] N \partial_{r_i} \left([N^{-1} \partial_{r_{j_n}}^{b_{j_n}^n} \log \phi_{\chi, N}^I(rN)] \cdots [N^{-1} \partial_{r_{j_n}}^{b_{j_n}^n} \log \phi_{\chi, N}^I(rN)] \right),$$

where the final term converges to an analytic function of r^I by Definition 2.4. By the same reasoning as for (c) and (d), we obtain (e). Finally, Claim (f) follows from the fact that

$$\#\{\text{appearances of } \log \phi_{\chi, N}^I(s)\} = l_n = k_n - |A_n| - \sum_i (b_i^n - 1) = k_n - W_{T, L}.$$

For the inductive step, Claims (a) and (b) follow for the same reasons. For (c), we see that $N^{W_{T_t}, L} X_t^I(rN)$ takes the form

$$\begin{aligned} N^{W_{T_t}, L} X_t^I(rN) &= \text{Sym}_{K_{T_t}} \left[\prod_{a \in A_m} \frac{1}{r_{j_m} - r_a} \left[[N^{W_{i_1}, L} \partial_{r_{j_m}}^{c_m^1} Y_{i_1}^I(rN)] \cdots [N^{W_{i_k}, L} \partial_{r_{j_m}}^{c_m^k} Y_{i_k}^I(rN)] \right] \right. \\ &\quad \left. [\partial_{r_{j_m}}^{b_{j_m}^m} \log \phi_{\chi, N}^I(rN)] \cdots [\partial_{r_{j_m}}^{b_{j_m}^m} \log \phi_{\chi, N}^I(rN)] \right], \end{aligned}$$

where Y_i are terms with weights $W_{i, L}$ appearing in the expansion of $\frac{D_{k_m+1} \cdots D_{k_n} \phi_{\chi, N}(s)}{\phi_{\chi, N}(s)}$ and $c_m^i > 0$ are such that $c_m^1 + \cdots + c_m^k = c_m$. Claim (c) then follows by the inductive hypothesis, Definition 2.3, and the same argument as in the base case. For (d), we notice that

$$\begin{aligned} N^{W_{T_t}, C} X_t^I(rN) &= \text{Sym}_{K_{T_t}} \left[\prod_{a \in A_m} \frac{1}{r_{j_m} - r_a} \left[[N^{W_{i_1}, C+1} \partial_{r_{j_m}}^{c_m^1} Y_{i_1}^I(rN)] \cdots [N^{W_{i_k}, C+1} \partial_{r_{j_m}}^{c_m^k} Y_{i_k}^I(rN)] \right] \right. \\ &\quad \left. [\partial_{r_{j_m}}^{b_{j_m}^m} \log \phi_{\chi, N}^I(rN)] \cdots [\partial_{r_{j_m}}^{b_{j_m}^m} \log \phi_{\chi, N}^I(rN)] \right]. \end{aligned}$$

By the inductive hypothesis and the fact that CLT-appropriate measures are LLN-appropriate in Definition 2.4, all terms after $\text{Sym}_{K_{T_t}} \left[\prod_{a \in A_m} \frac{1}{r_{j_m} - r_a} \right]$ converge to an analytic function in r^I , hence so does the final term upon application of Lemma 2.9. For (e), if $i \notin K_{T_t}$, we notice that ∂_{r_i} and $\text{Sym}_{K_{T_t}}$ commute and we obtain

$$\begin{aligned} N^{W_{T_t}, C+1} \partial_{r_i} X_t^I(rN) &= \text{Sym}_{K_{T_t}} \left[\prod_{a \in A_m} \frac{1}{r_{j_m} - r_a} N \partial_{r_i} \left([N^{W_{i_1}, C+1} \partial_{r_{j_m}}^{c_m^1} Y_{i_1}^I(rN)] \cdots \right. \right. \\ &\quad \left. \left. [N^{W_{i_k}, C+1} \partial_{r_{j_m}}^{c_m^k} Y_{i_k}^I(rN)] [\partial_{r_{j_m}}^{b_{j_m}^m} \log \phi_{\chi, N}^I(rN)] \cdots [\partial_{r_{j_m}}^{b_{j_m}^m} \log \phi_{\chi, N}^I(rN)] \right) \right]. \end{aligned}$$

By a combination of the inductive hypothesis and Definition 2.4, we see that all terms after $\text{Sym}_{K_{T_t}} \left[\prod_{a \in A_m} \frac{1}{r_{j_m} - r_a} \right]$ converge to an analytic function in r^I , so Claim (e) follows from Lemma 2.9. Finally, Claim (f) now follows from our recursive definition of the LLN weight and the fact that $k_m = |A_m| + \sum_{i=1}^{l_m} b_i^m + c_m$. $\square$

Corollary 2.12. If a term $X_1 \cdots X_l$ corresponds to a forest with trees $\{T_1, \dots, T_l\}$ and index sets $K_{T_1}, \dots, K_{T_l}$, then for $t \in \{1, \dots, l\}$, $X_t^I(s)$ satisfies

- (a) if $d\mu_N$ is LLN-appropriate for χ_N , uniformly in s^I/N in a neighborhood of $V_\chi^{|I|}$, $m_1, \dots, m_k \geq 0$, and $i_1, \dots, i_k \in K_{T_t}$, we have that

$$\partial_{s_{i_1}}^{m_1} \cdots \partial_{s_{i_k}}^{m_k} X_t^I(s) = O\left(N^{-W_{T_t}, L - \sum_j m_j}\right);$$

- (b) if $d\mu_N$ is CLT-appropriate for χ_N , uniformly in s^I/N in a neighborhood of $V_\chi^{|I|}$, $m_1, \dots, m_k \geq 0$, and $i_1, \dots, i_k \in K_{T_t}$, we have that

$$\partial_{s_{i_1}}^{m_1} \dots \partial_{s_{i_k}}^{m_k} X_t^I(s) = O\left(N^{-W_{T_t, C} - \sum_j m_j - \mathbf{1}_{\{m_j > 0 \text{ for some } i_j \notin K_{T_t}\}}}\right).$$

Proof. Claim (a) follows from Lemma 2.11(c), and Claim (b) follows from Lemma 2.11(d) and (e). $\square$

2.5. Proof of Theorem 2.5. Define the expressions

$$\tilde{e}_{a,i}(s) := a!e_a\left(\frac{1}{s_i - s_\star}\right)_{\star \neq i} \quad \text{and} \quad \tilde{e}_a(u, s) := a!e_a\left(\frac{1}{u - s_\star}\right),$$

where e_a denotes the a^{th} elementary symmetric polynomial, and $(s_1, \dots, s_N)$ is either a collection of variables or numbers with $s_i \neq s_j$ for any $j \neq i$.

Lemma 2.13. For $u \in \mathbb{C} \setminus V_\chi$, we have that

$$N^{-a}\tilde{e}_a(u, \chi_N/N) = \Xi(u)^a + O(N^{-1}).$$

Proof. Define the normalized power sums by

$$\tilde{p}_k(u, s) := \frac{1}{N} \sum_{i=1}^N \frac{1}{(u - s_i)^k}.$$

Viewing $\tilde{p}_k(u, \chi_N/N)$ as a Riemann sum, we find that

$$(2.13) \quad \tilde{p}_k(u, \chi_N/N) = \int_{-\infty}^{\infty} \frac{1}{(u - x)^k} d\chi(x) + O(N^{-1}) = \begin{cases} \Xi(u) & k = 1 \\ O(1) & k > 1 \end{cases} + O(N^{-1}).$$

Now, by [Mac95, Example I.2.8], we have the determinant formula expressing elementary symmetric polynomials in terms of power sums

$$e_a = \frac{1}{a!} \begin{vmatrix} p_1 & 1 & 0 & \cdots & 0 \\ p_2 & p_1 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ p_{a-1} & p_{a-2} & \cdots & & a-1 \\ p_a & p_{a-1} & \cdots & & p_1 \end{vmatrix}.$$

If we apply this formula to compute $\tilde{e}_a(u, \chi_N/N)$, each factor of p_k has order N by (2.13). Therefore, the leading order contribution must come from the product of the diagonal entries, implying that

$$N^{-a}\tilde{e}_a(u, \chi_N/N) = \tilde{p}_1(u, \chi_N/N)^a(1 + O(N^{-1})) = \Xi(u)^a + O(N^{-1}). \quad \square$$

We are now ready to prove the law of large numbers. By Lemma 2.2, we have that

$$\frac{1}{N} \mathbb{E}[p_k(x)] = \frac{1}{N} \frac{D_k \phi_{\chi, N}(\chi_N)}{\phi_{\chi, N}(\chi_N)}.$$

Applying the graphical calculus of the previous section and noting the additional prefactor of N^{-1} , this is a linear combination of the evaluation at $s = \chi_N$ of terms of the form

$$N^{-|A_1|-1} \text{Sym}_{\{j_1\} \cup A_1} \left[\prod_{a \in A_1} \frac{1}{s_{j_1} - s_a} \right] [\partial_{j_1}^{b_1^1} \log \phi_{\chi, N}(s)] \cdots [\partial_{j_n}^{b_n^1} \log \phi_{\chi, N}(s)],$$

where $j_1 \notin A_1$, $b_h^1 > 0$, and $|A_1| + \sum_h b_h^1 = k$. Such a term corresponds to 1-vertex tree with index set $K = \{j_1\} \cup A_1$ and LLN weight $|A_1| + \sum_h (b_h^1 - 1)$, meaning by Corollary 2.11(c) that it has order $O(N^{-|A_1| - \sum_h (b_h^1 - 1)})$; there are $O(N^{|A_1|+1})$ such terms, meaning that their sum has order $O(N^{-\sum_h (b_h^1 - 1)})$. As a result, the terms with leading order contribution are those with $b_h^1 = 1$, implying that

$$\frac{1}{N} \mathbb{E}[p_k(x)] = \frac{1}{N} \sum_{b=0}^k \binom{k}{b} N^{b-k} \sum_{j=1}^N \tilde{e}_{k-b,j}(r) [N^{-1} \partial_{r_j} \log \phi_{\chi, N}(rN)]^b \Big|_{r=\chi_N/N} + O(N^{-1}).$$

Notice now that

$$\sum_{j=1}^N \tilde{e}_{k-b,j}(r) [N^{-1} \partial_{r_j} \log \phi_{\chi,N}(rN)]^b \Big|_{r=\chi_N/N}$$

is a linear combination with coefficients independent of N of $O(N^{k-b+1})$ terms of the form

$$\text{Sym}_{\{j\} \cup A} \left[\prod_{a \in A} \frac{1}{r_j - r_a} \right] [N^{-1} \partial_{r_j} \log \phi_{\chi,N}^I(rN)]^b \Big|_{r=\chi_N/N}$$

for $j \notin A$, $|A| = k - b$, and $I = \{j\} \cup A$. By LLN-appropriateness we see that

$$[N^{-1} \partial_{r_j} \log \phi_{\chi,N}^I(rN)]^b = \Psi(r_j)^b + o(1),$$

uniformly on $U^{|I|}$ for an open neighborhood $U \supset V_\chi$. By Lemma 2.9, this implies that

$$\text{Sym}_{\{j\} \cup A} \left[\prod_{a \in A} \frac{1}{r_j - r_a} \right] \left([N^{-1} \partial_{r_j} \log \phi_{\chi,N}^I(rN)]^b - \Psi(r_j)^b \right) \Big|_{r=\chi_N/N} = o(1)$$

uniformly on compact subsets of $U^{|I|}$. This implies that

$$\begin{aligned} \frac{1}{N} \sum_{b=0}^k \binom{k}{b} N^{b-k} \sum_{j=1}^N \tilde{e}_{k-b,j}(r) \left([N^{-1} \partial_{r_j} \log \phi_{\chi,N}(rN)]^b - \Psi(r_j)^b \right) \Big|_{r=\chi_N/N} \\ = \frac{1}{N} \sum_{b=0}^k \binom{k}{b} N^{b-k} \cdot o(N^{k-b+1}) = o(1). \end{aligned}$$

We conclude that

$$\begin{aligned} \frac{1}{N} \mathbb{E}[p_k(x)] &= \frac{1}{N} \sum_{b=0}^k N^{b-k} \binom{k}{b} \sum_{j=1}^N \tilde{e}_{k-b,j}(r) \left(\Psi(r)^b + o(1) \right) \Big|_{r=\chi_N/N} \\ &= \frac{1}{N} \sum_{b=0}^k \binom{k}{b} \oint \frac{du}{2\pi i} N^{b-k} \frac{\tilde{e}_{k-b+1}(u, \chi_N/N)}{k-b+1} \Psi(u)^b + o(1), \end{aligned}$$

where the u -contour encloses V_χ and lies within U . By Lemma 2.13, we find that

$$\begin{aligned} \frac{1}{N} \mathbb{E}[p_k(x)] &= \sum_{b=0}^k \binom{k}{b} \oint \frac{du}{2\pi i} \frac{\Xi(u)^{k-b+1}}{k-b+1} \Psi(u)^b + o(1) \\ &= \frac{1}{k+1} \sum_{b=0}^k \binom{k+1}{b} \oint \frac{du}{2\pi i} \Xi(u)^{k-b+1} \Psi(u)^b + o(1) \\ &= \frac{1}{k+1} \oint \frac{du}{2\pi i} \left(\Xi(u) + \Psi(u) \right)^{k+1} + o(1). \end{aligned}$$

It remains to check that $\lim_{N \rightarrow \infty} \frac{1}{N^2} \text{Cov}(p_k(x), p_k(x)) = 0$. For this, define

$$\begin{aligned} (2.14) \quad C_{N,k,l}(s) &:= \frac{D_k D_l \phi_{\chi,N}(s)}{\phi_{\chi,N}(s)} - \frac{D_l \phi_{\chi,N}(s)}{\phi_{\chi,N}(s)} \frac{D_l \phi_{\chi,N}(s)}{\phi_{\chi,N}(s)} \\ &= \sum_{a=0}^k \sum_{l=0}^b \binom{k}{a} \binom{l}{b} \sum_{c=1}^a \sum_{i=1}^N \binom{a}{c} \tilde{e}_{k-a,i}(s) \partial_{s_i}^c \left[\sum_{j=1}^N \tilde{e}_{l-b,j}(s) \frac{\partial_{s_j}^b \phi_{\chi,N}(s)}{\phi_{\chi,N}(s)} \right] \frac{\partial_{s_i}^{a-c} \phi_{\chi,N}(s)}{\phi_{\chi,N}(s)} \end{aligned}$$

and notice by Lemma 2.2 that

$$\frac{1}{N^2} \text{Cov}(p_k(x), p_k(x)) = \frac{1}{N^2} C_{N,k,k}(\chi_N).$$

Choose indices i, j and sets A_1, A_2 with $i \notin A_1$ and $j \notin A_2$. Defining $K_1 := \{i\} \cup \{j\} \cup A_1 \cup A_2$ and $K_2 := \{j\} \cup A_2$, we notice that $C_{N,k,k}(\chi_N)$ is a linear combination of the evaluation at χ_N of terms of the form

$$(2.15) \quad N^{-|A_1|-|A_2|} \text{Sym}_{K_1} \left[\prod_{* \in A_1} \frac{1}{s_i/N - s_*/N} \right] \prod_{h=1}^{t'} [\partial_{s_i}^{b'_h} \log \phi_{\chi,N}^{K_1}(s)] \\ \partial_{s_i}^c \left[\text{Sym}_{K_2} \left[\prod_{* \in A_2} \frac{1}{s_j/N - s_*/N} \right] \prod_{h=1}^t [\partial_{s_j}^{b_h} \log \phi_{\chi,N}^{K_1}(s)] \right]$$

for different choices of i, j, A_1, A_2 and $c, b_h, b'_h > 0$ with $|A_1| + c + \sum_h b'_h = k$, $|A_2| + \sum_h b_h = l$, and $c \geq 1$. Such terms are again linear combinations of the terms associated to the forest on $\{1, 2\}$ consisting of a 2-vertex tree with index set K_1 and LLN weight

$$|A_1| + |A_2| + \sum_h (b'_h - 1) + \sum_h (b_h - 1) + c,$$

so by Lemma 2.11(c), they have order

$$O\left(N^{-|A_1|-|A_2|-\sum_h (b'_h-1)-\sum_h (b_h-1)-c}\right).$$

There are $O(N^{|A_1|+|A_2|+1+\mathbf{1}_{i \notin K_2}})$ such terms, meaning that their sum has order

$$O\left(N^{-\sum_h (b'_h-1)-\sum_h (b_h-1)+(1-c)+\mathbf{1}_{i \notin K_2}}\right),$$

hence the terms with leading order contribution to $C_{N,k,k}(\chi_N)$ are those with $b'_h = 1$, $b_h = 1$, $c = 1$, and $i \notin K_2$. We conclude that $C_{N,k,k}(\chi_N) = O(N)$, hence

$$\frac{1}{N^2} \text{Cov}(p_k(x), p_k(x)) = \frac{1}{N^2} C_{N,k,k}(\chi_N) = O(N^{-1}),$$

which concludes the proof that $\lim_{N \rightarrow \infty} \frac{1}{N} p_k(x) = \lim_{N \rightarrow \infty} \mathbb{E}[p_k(x)]$ in probability. This immediately implies the corresponding convergence of empirical measures.

2.6. Proof of Theorem 2.6. We prove Theorem 2.6 in two steps. We first prove that $\text{Cov}(p_k(x), p_l(x))$ has the claimed value and then show that all higher cumulants vanish.

2.6.1. Computing the covariance. Notice that $\text{Cov}(p_k(x), p_l(x)) = C_{N,k,l}(\chi_N)$, where we recall the definition of $C_{N,k,l}(s)$ from (2.14). Applying the expansion from the previous section into terms of the form (2.15), we see that $C_{N,k,l}(\chi_N)$ is a linear combination of terms associated to the forest on $\{1, 2\}$ consisting of a 2-vertex tree with index sets $K_1 = \{i\} \cup A_1 \cup \{j\} \cup A_2$ and $K_2 = \{j\} \cup A_2$ and CLT weight

$$|A_1| + |A_2| + \sum_h (b'_h - 1) + \sum_h (b_h - 1) + c + \mathbf{1}_{i \notin K_2},$$

so by Lemma 2.11(d), they have order

$$O\left(N^{-|A_1|-|A_2|-\sum_h (b'_h-1)-\sum_h (b_h-1)-c-\mathbf{1}_{i \notin K_2}}\right).$$

There are $O(N^{|A_1|+|A_2|+1+\mathbf{1}_{i \notin K_2}})$ such terms, meaning that their sum has order

$$O(N^{-\sum_h (b'_h-1)-\sum_h (b_h-1)+(1-c)}).$$

Thus, the leading order terms in $C_{N,k,l}(\chi_N)$ are those with $b'_h = 1$, $b_h = 1$, and $c = 1$. To compute the covariance, we analyze these terms more precisely. Setting $I = K_1$, define

$$\Delta_N^{I,j}(r) := \frac{1}{N} \partial_{r_j} \log \phi_{\chi,N}^I(rN) - \Psi(r_j)$$

so that from Definition 2.4 we have

$$(2.16) \quad \partial_{r_j}^m \Delta_N^{I,j}(r) = o(1), \quad m \geq 0 \\ \partial_{r_i} \Delta_N^{I,j}(r) = \frac{1}{N} F^{(1,1)}(r_i, r_j) + o(N^{-1}), \quad i \neq j.$$

With this notation, the terms with $b'_h = 1$, $b_h = 1$, and $c = 1$ take the form

$$\begin{aligned} & N^{-|A_1|-|A_2|-1} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} \left[\text{Sym}_{K_2} \prod_{a \in A_2} \frac{1}{r_j - r_a} [N^{-1} \partial_{r_j} \log \phi_{\chi, N}^I(rN)]^{l-|A_2|} \right] [N^{-1} \partial_{r_i} \log \phi_{\chi, N}^I(rN)]^{k-|A_1|-1} \right] \\ &= N^{-|A_1|-|A_2|-1} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} \left[\text{Sym}_{K_2} \prod_{a \in A_2} \frac{1}{r_j - r_a} \left(\Psi(r_j) + \Delta_N^{I,j}(r) \right)^{l-|A_2|} \right] \left(\Psi(r_i) + \Delta_N^{I,i}(r) \right)^{k-|A_1|-1} \right] \\ &= N^{-|A_1|-|A_2|-1} \sum_{m=0}^{l-|A_2|} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} [f_{j, A_2}^m(r)] \left(\Psi(r_i) + \Delta_N^{I,i}(r) \right)^{k-|A_1|-1} \right] \end{aligned}$$

for some $i \notin A_1$ and $j \notin A_2$ and $f_{j, A_2}^m(r)$ defined for $0 \leq m \leq l - |A_2|$ by

$$f_{j, A_2}^m(r) := \text{Sym}_{K_2} \prod_{a \in A_2} \frac{1}{r_j - r_a} \binom{l-|A_2|}{m} \Psi(r_j)^{l-|A_2|-m} \Delta_N^{I,j}(r)^m.$$

We study the asymptotics of each summand separately. We see that each $f_{j, A_2}^m(r)$ converges as $N \rightarrow \infty$ to an analytic function of r^I , which is 0 unless $m = 0$. Therefore, if $i \in K_2$ and $m > 0$ the summand

$$N^{-|A_1|-|A_2|-1} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} [f_{j, A_2}^m(r)] \left(\Psi(r_i) + \Delta_N^{I,i}(r) \right)^{k-|A_1|-1} \right]$$

has analytic limit of order $o(N^{-|A_1|-|A_2|-1})$ by Lemma 2.9. Similarly if $m = 0$, for $g > 0$, the terms

$$(2.17) \quad N^{-|A_1|-|A_2|-1} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} [f_{j, A_2}^0(r)] \binom{k-|A_1|-1}{g} \Psi(r_i)^{k-|A_1|-1-g} \Delta_N^{I,i}(r)^g \right]$$

resulting from the binomial expansion of $\left(\Psi(r_i) + \Delta_N^{I,i}(r) \right)^{k-|A_1|-1}$ have analytic limit of order $o(N^{-|A_1|-|A_2|-1})$ by Lemma 2.9. There are $O(N^{|K_2|}) = O(N^{|A_1|+|A_2|+1})$ of each of these type of terms, so the resulting terms do not contribute to the limit unless $m = 0$ and $g = 0$.

Now, if $i \notin K_2$, for $m \geq 1$ we have that

$$\partial_{r_i} f_{j, A_2}^m(r) = \frac{1}{N} \text{Sym}_{K_2} \prod_{a \in A_2} \frac{1}{r_j - r_a} \binom{l-|A_2|}{m} \Psi(r_j)^{l-|A_2|-m} m \Delta_N^{I,j}(r)^{m-1} F^{(1,1)}(r_i, r_j) + o(N^{-1}),$$

meaning that $N \partial_{r_i} f_{j, A_2}^m(r)$ converges as $N \rightarrow \infty$ to an analytic function of r^I which is 0 unless $m = 1$. Therefore, if $i \notin K_1$, for $m > 1$ the summand

$$N^{-|A_1|-|A_2|-1} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} [f_{j, A_2}^m(r)] \left(\Psi(r_i) + \Delta_N^{I,i}(r) \right)^{k-|A_1|-1} \right]$$

has analytic limit of order $o(N^{-|A_1|-|A_2|-2})$ by Lemma 2.9. Similarly, if $m = 1$, for $g > 0$ the terms

$$(2.18) \quad N^{-|A_1|-|A_2|-1} \text{Sym}_{K_1} \left[\prod_{a \in A_1} \frac{1}{r_i - r_a} \partial_{r_i} [f_{j, A_2}^1(r)] \binom{k-|A_1|-1}{g} \Psi(r_i)^{k-|A_1|-1-g} \Delta_N^{I,i}(r)^g \right]$$

have analytic limit of order $o(N^{-|A_1|-|A_2|-2})$ by Lemma 2.9. Since there are $O(N^{|K_2|}) = O(N^{|A_1|+|A_2|+2})$ such terms, their total contribution is $o(1)$ unless $m = 1$ and $g = 0$.

We conclude that the leading order terms in $C_{N, k, l}(\chi_N)$ correspond to those with $m = 0$, $g = 0$, and $i \in K_2$ in (2.17) or $m = 1$, $g = 0$, and $i \notin K_2$ in (2.18), with their coefficients given by their coefficients in

the direct expansions $X_1(\chi_N/N)$ and $X_2(\chi_N/N)$ for

$$X_1(r) = \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} \sum_{i=1}^N a N^{a-k+b-l-1} \tilde{e}_{k-a,i}(r) \partial_{r_i} \left[\sum_{j=1}^N \tilde{e}_{l-b,j}(r) \Psi(r_j)^b \right] \Psi(r_i)^{a-1}$$

$$X_2(r) = \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} \sum_{i=1}^N a N^{a-k+b-l-2} \tilde{e}_{k-a,i}(r) \sum_{j \neq i} \sum_{\substack{A_2: i, j \notin A_2 \\ |A_2|=l-b}} (l-b)! \prod_{a \in A_2} \frac{1}{r_j - r_a} b \Psi(r_j)^{b-1} F^{(1,1)}(r_i, r_j) \Psi(r_i)^{a-1},$$

where $X_1(\chi_N/N)$ corresponds to the terms with $m = 0$, and $X_2(\chi_N/N)$ corresponds to the terms with $m = 1$. Defining $\hat{\chi}_N := \chi_N/N$, we see that

$$\begin{aligned} X_1(\hat{\chi}_N) &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} a N^{a-k+b-l-1} \sum_{i=1}^N \tilde{e}_{k-a,i}(r) \partial_{r_i} \left[\oint \frac{du}{2\pi i} \frac{\tilde{e}_{l-b+1}(u, r)}{l-b+1} \Psi(u)^b \right] \Psi(r_i)^{a-1} \Big|_{r=\hat{\chi}_N} \\ &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} a N^{a-k+b-l-1} \sum_{i=1}^N \tilde{e}_{k-a,i}(r) \oint \frac{du}{2\pi i} \sum_{p=0}^{l-b} \frac{(-1)^p (l-b)_p}{(u-r_i)^{p+2}} \tilde{e}_{l-b-p}(u, r) \Psi(u)^b \Psi(r_i)^{a-1} \Big|_{r=\hat{\chi}_N} \\ &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} a N^{a-k+b-l-1} \oint \oint \frac{dw}{2\pi i} \frac{du}{2\pi i} \frac{\tilde{e}_{k-a+1}(w, \hat{\chi}_N)}{k-a+1} \sum_{p=0}^{l-b} \frac{(-1)^p (l-b)_p}{(u-w)^{p+2}} \tilde{e}_{l-b-p}(u, \hat{\chi}_N) \Psi(u)^b \Psi(w)^{a-1}, \end{aligned}$$

where both the u and w -contours enclose V_χ and lie in U , and the u -contour is contained inside the w -contour.

Applying Lemma 2.13 yields

$$\begin{aligned} X_1(\hat{\chi}_N) &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} a N^{-p} \oint \oint \frac{dw}{2\pi i} \frac{du}{2\pi i} \frac{\Xi(w)^{k-a+1}}{k-a+1} \sum_{p=0}^{l-b} \frac{(-1)^p (l-b)_p}{(u-w)^{p+2}} \Xi(u)^{l-b-p} \Psi(u)^b \Psi(w)^{a-1} + O(N^{-1}) \\ &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} a \oint \oint \frac{dw}{2\pi i} \frac{du}{2\pi i} \frac{\Xi(w)^{k-a+1}}{k-a+1} \Xi(u)^{l-b} \frac{\Psi(u)^b \Psi(w)^{a-1}}{(u-w)^2} + O(N^{-1}). \end{aligned}$$

Finally, adding a (vanishing) term with $a = k+1$ in the second from last equality, we obtain

$$\begin{aligned} X_1(\hat{\chi}_N) &= \sum_{a=1}^{k+1} \sum_{b=0}^l \binom{k}{a-1} \binom{l}{b} \oint \oint \frac{dw}{2\pi i} \frac{du}{2\pi i} \Xi(w)^{k-a+1} \Xi(u)^{l-b} \frac{\Psi(u)^b \Psi(w)^{a-1}}{(u-w)^2} + O(N^{-1}) \\ &= \oint \oint \frac{dw}{2\pi i} \frac{du}{2\pi i} \left(\Xi(w) + \Psi(w) \right)^k \left(\Xi(u) + \Psi(u) \right)^l \frac{1}{(u-w)^2} + O(N^{-1}), \end{aligned}$$

where both the u and w -contours enclose V_χ and lie in U , and the u -contour is contained inside the w -contour.

Notice now that

$$\begin{aligned} X_2(\hat{\chi}_N) &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} \sum_{i=1}^N a b N^{a-k+b-l-2} \tilde{e}_{k-a,i}(r) \\ &\quad \oint \frac{du}{2\pi i} (l-b)! e_{l-b+1} \left(\frac{1}{u-r_*} \right)_{* \neq i} \Psi(u)^{b-1} F^{(1,1)}(r_i, u) \Psi(r_i)^{a-1} \Big|_{r=\hat{\chi}_N} \\ &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} \sum_{i=1}^N a b N^{a-k+b-l-2} \tilde{e}_{k-a,i}(r) \\ &\quad \oint \frac{du}{2\pi i} \sum_{p=0}^{l-b+1} (-1)^p \frac{(l-b+1)_p}{(l-b+1)(u-r_i)^p} \tilde{e}_{l-b+1-p}(u, r) \Psi(u)^{b-1} F^{(1,1)}(r_i, u) \Psi(r_i)^{a-1} \Big|_{r=\hat{\chi}_N/N} \\ &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} a b N^{a-k+b-l-2} \sum_{p=0}^{l-b+1} (-1)^p \\ &\quad \oint \oint \frac{du}{2\pi i} \frac{dw}{2\pi i} \frac{\tilde{e}_{k-a+1}(w, \hat{\chi}_N)}{k-a+1} \frac{(l-b+1)_p}{(l-b+1)(u-w)^p} \tilde{e}_{l-b+1-p}(u, \hat{\chi}_N) \Psi(u)^{b-1} F^{(1,1)}(w, u) \Psi(w)^{a-1}, \end{aligned}$$

where both the u and w -contours enclose V_χ and lie in U , and the u -contour is contained inside the w -contour. Applying Lemma 2.13, we find that

$$\begin{aligned} X_2(\hat{\chi}_N) &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} ab N^{-p} \sum_{p=0}^{l-b+1} (-1)^p \oint \oint \frac{du}{2\pi i} \frac{dw}{2\pi i} \frac{\Xi(w)^{k-a+1}}{k-a+1} \\ &\quad \frac{(l-b+1)_p}{(l-b+1)(u-w)^p} \Xi(u)^{l-b+1-p} \Psi(u)^{b-1} F^{(1,1)}(w, u) \Psi(w)^{a-1} + O(N^{-1}) \\ &= \sum_{a=0}^k \sum_{b=0}^l \binom{k}{a} \binom{l}{b} ab \oint \oint \frac{du}{2\pi i} \frac{dw}{2\pi i} \frac{\Xi(w)^{k-a+1}}{k-a+1} \frac{\Xi(u)^{l-b+1}}{l-b+1} \Psi(u)^{b-1} F^{(1,1)}(w, u) \Psi(w)^{a-1} + O(N^{-1}), \end{aligned}$$

where both the u and w -contours enclose V_χ and lie in U , and the u -contour is contained inside the w -contour. Adding (identically zero) terms with $a = k+1$ and $b = l+1$ in the last equality, we find that

$$\begin{aligned} X_2(\hat{\chi}_N) &= \sum_{a=1}^{k+1} \sum_{b=1}^{l+1} \binom{k}{a-1} \binom{l}{b-1} \oint \oint \frac{du}{2\pi i} \frac{dw}{2\pi i} \Xi(w)^{k-a+1} \Xi(u)^{l-b+1} \Psi(u)^{b-1} F^{(1,1)}(w, u) \Psi(w)^{a-1} + O(N^{-1}) \\ &= \oint \oint \left(\Xi(w) + \Psi(w) \right)^k \left(\Xi(u) + \Psi(u) \right)^l F^{(1,1)}(w, u) \frac{du}{2\pi i} \frac{dw}{2\pi i} + O(N^{-1}), \end{aligned}$$

where both the u and w -contours enclose V_χ and lie in U , and the u -contour is contained inside the w -contour. Putting these computations together, we find the covariance to be

$$\begin{aligned} \text{Cov}(p_k(x), p_l(x)) &= X_1(\hat{\chi}_N) + X_2(\hat{\chi}_N) + O(N^{-1}) \\ &= \oint \oint \frac{dw}{2\pi i} \frac{du}{2\pi i} \left(\Xi(u) + \Psi(u) \right)^l \left(\Xi(w) + \Psi(w) \right)^k \left(\frac{1}{(u-w)^2} + F^{(1,1)}(w, u) \right) + O(N^{-1}), \end{aligned}$$

where both the u and w -contours enclose V_χ and lie in U with the u -contour contained inside the w -contour.

2.6.2. Establishing Gaussianity. For any $k_1, \dots, k_n$, define the cumulant by

$$\mathcal{K}_{k_1, \dots, k_n} := \sum_{U_1 \sqcup \dots \sqcup U_t = \{1, \dots, N\}} (-1)^{t-1} (t-1)! \prod_{i=1}^t \mathbb{E} \left[\prod_{j \in U_i} p_{k_j}(x) \right].$$

We now check that $\mathcal{K}_{k_1, \dots, k_n} = o(1)$ for $n \geq 3$. By an analogue of [BG18b, Lemma 3.10], the only terms in $\mathcal{K}_{k_1, \dots, k_n}$ with non-zero contribution correspond to a forest composed of a single tree and hence $t = 1$ and $U_1 = \{1, \dots, N\}$. We now check that these terms are $o(1)$. By construction, if a forest has a single tree, its index set is

$$K = \{j_1\} \cup \dots \cup \{j_n\} \cup A_1 \cup \dots \cup A_n$$

and it has CLT weight

$$\sum_{m=1}^n |A_m| + \sum_{i,m} (b_i^m - 1) + \sum_m c_m + \#\{\text{single edges in } T\},$$

where we notice that

$$\sum_m c_m \geq \#\{\text{edges in } T\}.$$

On the other hand, the size of the index set is at most

$$1 + \sum_{m=1}^n |A_m| + \#\{\text{single edges in } T\}.$$

By Lemma 2.11(d), the total contribution of these $O(N^{1+\sum_{m=1}^n |A_m| + \#\{\text{single edges in } T\}})$ terms has order

$$O\left(N^{\# \{1 - \sum_{i,m} (b_i^m - 1) - \sum_m c_m\}}\right),$$

where we notice that

$$1 - \sum_{i,m} (b_i^m - 1) - \sum_m c_m \leq 1 - \#\{\text{edges in } T\} \leq 2 - n.$$

This implies that the total contribution to the cumulant is $O(N^{2-n})$, which vanishes as desired for $n \geq 3$.

2.7. Central limit theorems for growing measures. In this section, we adapt our theorems to a case where the measures grow in a specific manner. Choose $M = \omega(1)$, where we recall that $f = \omega(g)$ if $\lim_{N \rightarrow \infty} f/g = \infty$, so that $M, N \rightarrow \infty$ simultaneously. As before, let $d\mu_N$ be a sequence of measures with multivariate Bessel generating function $\phi_{\chi, N}(s)$. Suppose that there exist a sequence of smooth measures $d\lambda'_N$ with multivariate Bessel generating function $\phi_{\chi, N}(s)^M$, and let $d\lambda_N$ be the pushforward of $d\lambda'_N$ under the map $\lambda' \mapsto \frac{1}{M}\lambda'$. In this setting, we have the following extensions of Theorem 2.5 and Theorem 2.6.

Theorem 2.14. If $d\mu_N$ is LLN-appropriate for χ_N , then if x is distributed according to $d\lambda_N$, we have the following convergence in probability

$$\lim_{N \rightarrow \infty} \frac{1}{N} p_k(x) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(x)] = \mathbf{p}'_k := \oint \Xi(u) \Psi(u)^k \frac{du}{2\pi \mathbf{i}},$$

where the u -contour encloses V_χ and lies within U . In addition, the random measures $\frac{1}{N} \sum_{i=1}^N \delta_{x_i}$ with x distributed according to $d\lambda_N$ converge in probability to a deterministic compactly supported measure $d\lambda$ with $\int x^k d\lambda(x) = \mathbf{p}'_k$.

Proof. Define $\Phi_{\chi, N}(s) := \phi_{\chi, N}(s)^M$. As in our previous analysis, we have that

$$\mathbb{E}[p_{k_1}(x) \cdots p_{k_n}(x)] = M^{-k_1 - \cdots - k_n} \frac{D_{k_1} \cdots D_{k_n} \Phi_{\chi, N}(\chi_N)}{\Phi_{\chi, N}(\chi_N)},$$

which may be written as a linear combination of terms encoded by the graphical calculus of Section 2.4. The coefficients of this linear combination are independent of N and are monomials in M . By Lemma 2.11(f), the exponent in M is bounded above by $-W_L$, where W_L is the LLN weight of the term. Since $M = \omega(1)$, the leading order terms in our situation are the subset of the leading order terms of the case $M = 1$ corresponding to the terms whose coefficient has the maximal power of M . In particular, this implies by the proof of Theorem 2.5 that

$$\begin{aligned} \frac{1}{N} \mathbb{E}[p_k(x)] &= \frac{M^{-k}}{k+1} \oint \left(\Xi(u) + M \Psi(u) \right)^{k+1} \frac{du}{2\pi \mathbf{i}} + O(N^{-1}) \\ &= \oint \Xi(u) \Psi(u)^k \frac{du}{2\pi \mathbf{i}} + O(M^{-1} + N^{-1}), \end{aligned}$$

where the contour encloses V_χ and lies within a neighborhood U of V_χ , and we notice that

$$\oint \Psi(u)^{k+1} \frac{du}{2\pi \mathbf{i}}$$

vanishes because Ψ is holomorphic on U . Similarly, we have from the proof of Theorem 2.5 that $\frac{1}{N^2} \text{Cov}(p_k(x), p_k(x)) = O(N^{-1})$, which yields the desired. $\square$

Theorem 2.15. If the measures $d\mu_N$ are CLT-appropriate for χ_N , then if x is distributed according to $d\lambda_N$, the collection of random variables

$$\{M^{1/2}(p_k(x) - \mathbb{E}[p_k(x)])\}_{k \in \mathbb{N}}$$

converges in the sense of moments to the Gaussian vector with zero mean and covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}(M^{1/2} p_k(x), M^{1/2} p_l(x)) \\ = kl \oint \oint \Xi(u) \Xi(w) \Psi(u)^{k-1} \Psi(w)^{l-1} F^{(1,1)}(u, w) \frac{du}{2\pi \mathbf{i}} \frac{dw}{2\pi \mathbf{i}} + kl \oint \Xi(u) \Psi(u)^{k+l-2} \Psi'(u) \frac{du}{2\pi \mathbf{i}}, \end{aligned}$$

where the u and w -contours enclose V_χ and lie within U , and the u -contour is contained inside the w -contour.

Proof. As in the proof of Theorem 2.6, we analyze for $n \geq 2$ the terms which appear in the expansion of the cumulant $\mathcal{K}_{k_1, \dots, k_n}$ of $p_{k_1}(x), \dots, p_{k_n}(x)$ as encoded in the graphical calculus of Section 2.4. In particular, each term with non-zero contribution corresponds to a single tree T on n vertices; such terms have LLN weight

$$(2.19) \quad W_L = \sum_{m=1}^n |A_m| + \sum_{m=1}^n \sum_i (b_i^m - 1) + \sum_{m=1}^n c_m \geq \#\{\text{edges in } T\} = n - 1.$$

By Lemma 2.11(f), this implies that the exponent of M in each term with non-zero contribution to an n^{th} mixed cumulant is at most $1 - n$.

We now analyze the covariance and the higher cumulants separately. For the higher cumulants, by the proof of Theorem 2.6 and what we just showed, any n^{th} cumulant of $\{p_k(x)\}$ has order at most $O(N^{2-n}M^{1-n})$, meaning that any n^{th} cumulant of $\{M^{1/2}(p_k(x) - \mathbb{E}[p_k(x)])\}$ has order at most $O(N^{2-n}M^{1-n/2}) = o(1)$ for $n \geq 3$.

For the covariance, we again analyze terms in the expansion as in the proof of Theorem 2.6. Notice that the coefficients of each term are simply multiplied by a power of M equal to the number of appearances of a (derivative) of $\log \phi_{\chi, N}^I(s)$ which appears in it. As we just argued, this power is non-positive, which means that any leading order term in this expansion comes from a leading order term in the case $M = 1$ and that all other terms have order at most $O(N^{-1})$. This implies that

$$(2.20) \quad \text{Cov}\left(M^{1/2}p_k(x), M^{1/2}p_l(x)\right) = M^{-k-l+1} \oint \oint \left(\Xi(u) + M\Psi(u)\right)^l \left(\Xi(w) + M\Psi(w)\right)^k \left(\frac{1}{(u-w)^2} + MF^{(1,1)}(w, u)\right) \frac{dw}{2\pi i} \frac{du}{2\pi i} + O(N^{-1}),$$

where the u and w -contours enclose V_χ and lie in U , and the u -contour is contained inside the w -contour. Observe now that for $1 \leq a \leq l$ and $1 \leq b \leq k$ we have

$$\begin{aligned} \oint \oint \Psi(u)^l \Psi(w)^k F^{(1,1)}(w, u) \frac{dw}{2\pi i} \frac{du}{2\pi i} &= 0 \\ \oint \oint \Xi(u)^a \Psi(u)^{l-a} \Psi(w)^k F^{(1,1)}(w, u) \frac{dw}{2\pi i} \frac{du}{2\pi i} &= 0 \\ \oint \oint \Xi(w)^b \Psi(u)^l \Psi(w)^{k-b} F^{(1,1)}(w, u) \frac{dw}{2\pi i} \frac{du}{2\pi i} &= 0 \\ \oint \oint \Psi(u)^l \Psi(w)^k \frac{1}{(u-w)^2} \frac{dw}{2\pi i} \frac{du}{2\pi i} &= 0 \\ \oint \oint \Xi(w) \Psi(u)^l \Psi(w)^{k-1} \frac{1}{(u-w)^2} \frac{du}{2\pi i} \frac{dw}{2\pi i} &= 0, \end{aligned}$$

where the first, third, fourth, and fifth integrals vanish because the integrand is holomorphic inside the u -contour, and the second integral vanishes because the integrand is holomorphic inside the w -contour. We conclude that

$$\begin{aligned} \text{Cov}\left(M^{1/2}p_k(x), M^{1/2}p_l(x)\right) &= kl \oint \oint \Xi(u) \Xi(w) \Psi(u)^{l-1} \Psi(w)^{k-1} F^{(1,1)}(w, u) \frac{du}{2\pi i} \frac{dw}{2\pi i} \\ &\quad + l \oint \oint \Xi(u) \Psi(u)^{l-1} \Psi(w)^k \frac{1}{(u-w)^2} \frac{du}{2\pi i} \frac{dw}{2\pi i} + O(N^{-1}). \end{aligned}$$

Computing the w -integral as the residue at $w = u$ in the second integral yields the claim. $\square$

3. ASYMPTOTICS OF MULTIVARIATE BESSEL FUNCTIONS

3.1. Statement of the results. Recall that $\rho = (N-1, \dots, 0)$. In this section, we study the asymptotics of the normalized multivariate Bessel function

$$\frac{\mathcal{B}(\mu, \lambda)}{\mathcal{B}(\rho, \lambda)}$$

when μ deviates from ρ in only k coordinates. In this case, we may parametrize μ as

$$(3.1) \quad \mu := (a_1, a_2, \dots, a_k, N-1, \dots, \widehat{b}_1, \dots, \widehat{b}_2, \dots, \widehat{b}_k, \dots, 0).$$

for some $a_1, \dots, a_k$ and $b_1 > \dots > b_k \in \{0, \dots, N-1\}$, where $\widehat{b}_i$ denotes that the coordinate b_i is missing. We obtain asymptotics for these μ from asymptotics for

$$\mu_{a,b} := (a, N-1, \dots, \widehat{b}, \dots, 0)$$

differing from ρ in a single coordinate. Our first two results of this section relate these two cases and give a new contour integral expression (3.2) for the normalized multivariate Bessel function when $\mu = \mu_{a,b}$ which will be crucial to our asymptotic analysis.

Proposition 3.1. For two k -tuples $a_1, \dots, a_k \in \mathbb{C}$ and $b_1 > \dots > b_k \in \{0, \dots, N-1\}$, we have that

$$\frac{\mathcal{B}(\mu, \lambda)}{\mathcal{B}(\rho, \lambda)} = (-1)^{\frac{k(k-1)}{2}} \frac{\prod_{m \neq l} (a_m - b_l)}{\Delta(a_m) \Delta(b_m)} \det \left(\frac{a_i - b_j}{a_i - b_j} \frac{\mathcal{B}(\mu_{a_i, b_j}, \lambda)}{\mathcal{B}(\rho, \lambda)} \right)_{i,j=1}^k.$$

Theorem 3.2. For any $a \in \mathbb{C} \setminus \{N-1, \dots, 0\}$ and $b \in \{N-1, \dots, 0\}$, we have that

$$(3.2) \quad \frac{\mathcal{B}(\mu_{a,b}, \lambda)}{\mathcal{B}(\rho, \lambda)} = (-1)^{N-b+1} \frac{\Gamma(N-b)\Gamma(b+1)\Gamma(a-N+1)(a-b)}{\Gamma(a+1)} \oint_{\{e^{\lambda_i}\}} \frac{dz}{2\pi i} \oint_{\{0,z\}} \frac{dw}{2\pi i} \cdot \frac{z^a w^{-b-1}}{z-w} \cdot \prod_{i=1}^N \frac{w - e^{\lambda_i}}{z - e^{\lambda_i}},$$

where the z -contour encloses the poles at e^{λ_i} and avoids the negative real axis and the w -contour encloses both 0 and the z -contour.

Remark. Proposition 3.1 and Theorem 3.2 are parallel to [GP15, Theorem 3.7] and [GP15, Theorem 3.8]. We note that Proposition 3.1 does not involve differentiation and thus takes a simpler form than [GP15, Theorem 3.7]; in particular checking that the expression of Proposition 3.1 equals 1 when $\mu = \rho$ is straightforward, while the corresponding check in [GP15, Theorem 3.7] is rather delicate. On the other hand, Theorem 3.2 is more complicated than [GP15, Theorem 3.8], as a single contour integral is replaced by a double contour integral.

The final result of this section, Theorem 3.4 gives asymptotics under certain asymptotic assumptions on λ in the case where finitely many b_i are perturbed. We first specify these asymptotic assumptions. For a sequence $\lambda^1, \lambda^2, \dots$ where λ^N has length N , define $d\rho_N$ and $d\tilde{\rho}_N$ by

$$d\rho_N := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^N} \quad d\tilde{\rho}_N := \frac{1}{N} \sum_{i=1}^N \delta_{e^{\lambda_i^N}}.$$

as the empirical measure of λ^N and its pushforward under the exponential map. We make the following assumption on $d\rho_N$.

Assumption 3.3. The measures $d\rho_N$ have support contained within a fixed finite interval I and converge weakly to a compactly supported measure $d\rho$.

For $\tilde{b}_1 > \dots > \tilde{b}_k \in \frac{1}{N}\mathbb{Z}$ with $\tilde{b}_i N \in \{0, \dots, N-1\}$, define the N -tuple

$$(3.3) \quad \mu_{\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k} := (\tilde{a}_1 N, \tilde{a}_2 N, \dots, \tilde{a}_k N, N-1, \dots, \widehat{\tilde{b}_1 N}, \dots, \widehat{\tilde{b}_2 N}, \dots, \widehat{\tilde{b}_k N}, \dots, 0).$$

We will study the $N \rightarrow \infty$ asymptotics of the quantity

$$\tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) := \frac{\mathcal{B}(\mu_{\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k}, \lambda^N)}{\mathcal{B}(\rho, \lambda^N)}.$$

Our answer will be expressed in terms of the moment generating function of the exponential moments and its S -transform, defined by

$$M_{\tilde{\rho}_N}(z) := \int \frac{e^x z}{1 - e^x z} d\tilde{\rho}_N(x) \quad S_{\tilde{\rho}_N}(z) := \frac{1+z}{z} M_{\tilde{\rho}_N}^{-1}(z),$$

where by $M_{\tilde{\rho}_N}^{-1}$ we mean the functional inverse of $M_{\tilde{\rho}_N}$ as a (uniformly convergent) power series in z . We also consider their limiting versions defined by

$$M_{\tilde{\rho}}(z) := \int \frac{e^x z}{1 - e^x z} d\rho(x) \quad S_{\tilde{\rho}}(z) := \frac{1+z}{z} M_{\tilde{\rho}}^{-1}(z).$$

Some analytic properties of $M_{\tilde{\rho}_N}(z)$, $M_{\tilde{\rho}}(z)$, $S_{\tilde{\rho}_N}(z)$, and $S_{\tilde{\rho}}(z)$ are discussed in Lemma 3.10 below. Define finally the function

$$(3.4) \quad \tilde{\Psi}_{\rho_N}(c) := c \log S_{\tilde{\rho}_N}(c-1) + \int \log \left((1-c)(-M_{\tilde{\rho}_N}^{-1}(c-1)^{-1} + e^s) \right) d\rho_N(s),$$

which is holomorphic on a neighborhood of $[0, 1]$ by Lemmas 3.10 and 3.11 and the fact that

$$\tilde{\Psi}_{\rho_N}(c) = c \log S_{\tilde{\rho}_N}(c-1) + \int \log \left(c S_{\tilde{\rho}_N}(c-1)^{-1} + (1-c)e^s \right) d\rho_N(s).$$

Theorem 3.4. If λ^N satisfies Assumption 3.3, for $\tilde{b}_1 > \dots > \tilde{b}_k \in \frac{1}{N}\mathbb{Z}$ with $\tilde{b}_i N \in \{0, \dots, N-1\}$, as a function of $\tilde{a}_1, \dots, \tilde{a}_k$, the quantity $\tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k)$ admits a holomorphic extension to an open neighborhood U^k of $[0, 1]^k$ such that uniformly on compact subsets and uniformly in $\tilde{b}_1, \dots, \tilde{b}_k$ we have

$$\begin{aligned} \tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) &= \frac{\prod_{m \neq l} (\tilde{a}_m - \tilde{b}_l)}{\Delta(\tilde{a})\Delta(\tilde{b})} \frac{\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1))\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))}{\prod_{i,j} [M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1)]} \\ &\quad \prod_{i=1}^k \left[\frac{(\tilde{a}_i - \tilde{b}_i)}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1))M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))}} \frac{\sqrt{S_{\tilde{\rho}_N}(\tilde{b}_i - 1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{b}_i)}}{\sqrt{S_{\tilde{\rho}_N}(\tilde{a}_i - 1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{a}_i)}} \right] [1 + o(1)]. \end{aligned}$$

Lemma 3.5. The expression

$$\begin{aligned} &\frac{\prod_{m \neq l} (\tilde{a}_m - \tilde{b}_l)}{\Delta(\tilde{a})\Delta(\tilde{b})} \frac{\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1))\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))}{\prod_{i,j} [M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1)]} \\ &\quad \prod_{i=1}^k \left[\frac{(\tilde{a}_i - \tilde{b}_i)}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1))M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))}} \frac{\sqrt{S_{\tilde{\rho}_N}(\tilde{b}_i - 1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{b}_i)}}{\sqrt{S_{\tilde{\rho}_N}(\tilde{a}_i - 1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{a}_i)}} \right] \end{aligned}$$

in Theorem 3.4 is holomorphic in $\tilde{a}_i, \dots, \tilde{a}_k$ on an open complex neighborhood U^k of $[0, 1]^k$.

Proof. Note that

$$\frac{\prod_{m \neq l} (\tilde{a}_m - \tilde{b}_l)}{\Delta(\tilde{a})\Delta(\tilde{b})} \frac{\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1))\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1))}{\prod_{i \neq j} (M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1))}$$

has no poles in codimension 1 as a function of $\tilde{a}_1, \dots, \tilde{a}_k, \tilde{b}_1, \dots, \tilde{b}_k$ on U^{2k} , hence is holomorphic by Riemann's second extension theorem (see [GR84, Theorem 7.1.2]). In addition, by Lemma 3.10, $M_{\tilde{\rho}_N}^{-1}(u-1)$ is meromorphic on a neighborhood $U \supset [0, 1]$ uniformly in N and avoids 0 and ∞ aside from a simple zero at 1 and a simple pole at 0. Noting that $M_{\tilde{\rho}_N}^{-1}(u-1) \in [0, -\infty)$ for $u \in [0, 1]$ and that

$$(3.5) \quad M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(u-1)) = \int \frac{e^s}{(1 - e^s M_{\tilde{\rho}_N}^{-1}(u-1))^2} d\rho_N(s),$$

on a neighborhood of U^2 the function

$$\frac{(a-b)^2}{(M_{\tilde{\rho}_N}^{-1}(a-1) - M_{\tilde{\rho}_N}^{-1}(b-1))^2 M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(a-1))M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(b-1))}$$

has no zeros and has poles contained in the union of hyperplanes $\{a=b\} \cup \{a=0\} \cup \{a=1\} \cup \{b=0\} \cup \{b=1\}$. By (3.5), we may easily check that it has no poles in codimension 1 and hence no poles on all of U^2 . Similarly, $S_{\tilde{\rho}_N}(u-1)$ is holomorphic and non-vanishing on some U . Therefore, for a small enough neighborhood U of $[0, 1]$, we may choose branches of the functions

$$\frac{a-b}{M_{\tilde{\rho}_N}^{-1}(a-1) - M_{\tilde{\rho}_N}^{-1}(b-1)} \frac{1}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(a-1))M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(b-1))}}$$

and $\sqrt{S_{\tilde{\rho}_N}(u-1)}$ so that they are holomorphic on U^2 and U , respectively. Combining these with the fact that $\tilde{\Psi}_{\tilde{\rho}_N}(u)$ is holomorphic shows that the expression is holomorphic on U^k , as needed. $\square$

The remainder of this section is devoted to the proofs of Proposition 3.1, Theorem 3.2, and Theorem 3.4.

3.2. Multivariate Bessel functions and hook Schur functions. In this section we derive some properties of multivariate Bessel functions through their connection to certain symmetric functions called Schur functions. Recall that a signature $\mu = (\mu_1, \dots, \mu_N)$ is an N -tuple of integers such that $\mu_1 \geq \dots \geq \mu_N$. For such a signature, we may define the Schur function $s_\mu(x)$ by

$$(3.6) \quad s_\mu(x) := \frac{\det(x_i^{\mu_j + N - j})_{i,j=1}^N}{\prod_{1 \leq i < j \leq N} (x_i - x_j)}$$

as a function on $x_1, \dots, x_N$. By the Weyl character formula, it is the character of the representation of $U(N)$ with highest weight μ . Recalling that $\rho = (N - 1, \dots, 0)$, if $\mu - \rho$ is a signature, the multivariate Bessel function may be expressed in terms of Schur functions as

$$(3.7) \quad \mathcal{B}(\mu, \lambda) = \frac{\Delta(e^\lambda) \Delta(\rho)}{\Delta(\lambda) \Delta(\mu)} s_{\mu - \rho}(e^\lambda).$$

We will consider cases where $\mu - \rho$ is a partition, meaning that $\mu_N - \rho_N \geq 0$.

For a Young diagram of a partition λ with k boxes on its main diagonal, we may represent λ in Frobenius notation

$$\lambda = (\alpha_1, \dots, \alpha_k \mid \beta_1, \dots, \beta_k),$$

where $\alpha_i = \lambda_i - i$, $\beta_i = \lambda'_i - i$, and λ' denotes the dual (transposed) partition to λ . If $a_1 - N + 1 > \dots > a_k - N + k > N$, then for the μ defined in (3.1) we see that

$$\mu - \rho = (a_1 - N + 1, a_2 - N + 2, \dots, a_k - N + k, k, \dots, k, k - 1, \dots, k - 1, \dots, 1, \dots, 1, 0, \dots, 0),$$

where k appears $N - b_1 - 1$ times and $k - i$ appears $b_i - b_{i+1} - 1$ times. In Frobenius notation, we have that

$$\mu - \rho = (a_1 - N, \dots, a_k - N \mid N - 1 - b_k, \dots, N - 1 - b_1).$$

The following lemma expresses a Schur function corresponding to a partition with k hooks in terms of individual hook Schur functions $s_{\alpha|\beta}(x)$, which will allow us to prove Proposition 3.1.

Lemma 3.6 ([Mac95, Example I.3.9]). If $\lambda = (\alpha_1, \dots, \alpha_k \mid \beta_1, \dots, \beta_k)$ in Frobenius notation, we have that

$$s_\lambda(x) = \det(s_{\alpha_i|\beta_j}(x))_{i,j=1}^k,$$

where

$$s_{\alpha|\beta}(x) = \sum_{i=0}^{\beta} (-1)^i h_{\alpha+1+i}(x) e_{\beta-i}(x),$$

where $h_k(x)$ and $e_k(x)$ denote the complete homogeneous and elementary symmetric polynomial, respectively.

In order to convert Lemma 3.6 into a similar formula for multivariate Bessel functions, we require the following version of Carlson's Theorem from [Car14] as stated in [Pil05]. This result allows us to show that identities of analytic functions which hold on certain specializations hold everywhere.

Theorem 3.7 (Carlson's Theorem). Suppose $f(z)$ is holomorphic in a neighborhood of $\{z \in \mathbb{C} \mid \Re z \geq 0\}$ and satisfies

- (a) exponential growth, meaning $f(z) = O(e^{C_1|z|})$ for $z \in \mathbb{C}$ for some $C_1 > 0$;
- (b) exponential growth of order at most π along the imaginary axis, meaning $f(iy) = O(e^{C_2|y|})$ for $y \in \mathbb{R}$ for some $0 < C_2 < \pi$.

If $f(n) = 0$ for all non-negative integers n , then $f(z) = 0$ identically.

Proof of Proposition 3.1. First, treating both sides of the desired equality as functions of $a_1 - N + 1, \dots, a_k - N + k$, by considering the explicit definition (2.1), we see they satisfy both exponential growth conditions of Carlson's theorem separately in each variable. Applying Carlson's theorem to the difference in each variable sequentially, it suffices to check this for integer a_i such that $a_1 - N + 1 > \dots > a_k - N + k > N$ so that $\mu - \rho$ is a partition. In this case, combining the relation (3.7) and Lemma 3.6, we see that

$$(3.8) \quad \frac{\mathcal{B}(\mu, \lambda)}{\mathcal{B}(\rho, \lambda)} = \frac{\Delta(\rho)}{\Delta(\mu)} \det \left(s_{a_i - N \mid N - 1 - b_{k+1-j}}(e^\lambda) \right)_{i,j=1}^k.$$

We see that

$$\frac{\Delta(\rho)}{\Delta(\mu)} = \frac{\Delta(b_m)}{\Delta(a_m)} \prod_{m=1}^k (-1)^{N-b_m+m} \prod_{i \notin \{b_1, \dots, b_k\}} \frac{b_m - i}{a_m - i} = \frac{\prod_{m \neq l} (a_m - b_l)}{\Delta(b_m) \Delta(a_m)} \prod_{m=1}^k (-1)^{N-b_m+1} \prod_{i \neq b_m} \frac{b_m - i}{a_m - i}.$$

Moving these factors into the determinant in (3.8) and noting that

$$\frac{\mathcal{B}(\mu_{a,b}, \lambda)}{\mathcal{B}(\rho, \lambda)} = \frac{\Delta(\rho)}{\Delta(\mu_{a,b})} s_{a-N|N-1-b}(e^\lambda) = (-1)^{N-b+1} \prod_{i \neq b} \frac{b-i}{a-i} s_{a-N|N-1-b}(e^\lambda),$$

we find that

$$\frac{\mathcal{B}(\mu, \lambda)}{\mathcal{B}(\rho, \lambda)} = \frac{\prod_{m \neq l} (a_m - b_l)}{\Delta(a_m) \Delta(b_m)} \det \left(\frac{a_i - b_i}{a_i - b_{k+1-j}} \frac{\mathcal{B}(\mu_{a_i, b_{k+1-j}}, \lambda)}{\mathcal{B}(\rho, \lambda)} \right)_{i,j=1}^k,$$

which yields the result after applying the reindexing $j \mapsto k+1-j$. $\square$

3.3. An integral formula for multivariate Bessel functions. In this section we prove Theorem 3.2 by combining index-variable duality for Schur functions and the following integral formula for hook Schur functions.

Proposition 3.8. We have that

$$s_{\alpha|\beta}(x) = \oint_{\{0\}} \oint_{\{0\}} \frac{dz dw}{(2\pi i)^2} (-1)^\beta z^{-\alpha-\beta-2} w^{-N-1} \frac{1-w^{\beta+1}}{1-w} \prod_{i=1}^N \frac{w - x_i z}{1 - x_i z},$$

where both contours are small circles around 0.

Proof. We recall that for $H(x, z) := \prod_{i=1}^N (1 - zx_i)^{-1}$, we have that

$$h_l(x) = \frac{1}{2\pi i} \oint_{\{0\}} \frac{H(x, z)}{z^{l+1}} dz.$$

We conclude by Lemma 3.6 that

$$\begin{aligned} s_{\alpha|\beta}(x) &= \oint_{\{0\}} \frac{dz}{2\pi i} \sum_{j=0}^{\beta} (-1)^j \frac{e_{\beta-j}(x)}{z^{\alpha+2+j}} H(x, z) \\ &= \oint_{\{0\}} \frac{dz}{2\pi i} H(x, z) z^{-\alpha-\beta-2} (-1)^\beta \sum_{j=0}^{\beta} (-1)^{\beta-j} e_{\beta-j} z^{\beta-j} \\ &= \oint_{\{0\}} \oint_{\{0\}} \frac{dz dw}{(2\pi i)^2} H(x, z) z^{-\alpha-\beta-2} (-1)^\beta \prod_{i=1}^N (w - x_i z) (w^{-N-1} + \dots + w^{-N-1+\beta}) \\ &= \oint_{\{0\}} \oint_{\{0\}} \frac{dz dw}{(2\pi i)^2} (-1)^\beta z^{-\alpha-\beta-2} w^{-N-1} \frac{1-w^{\beta+1}}{1-w} \prod_{i=1}^N \frac{w - x_i z}{1 - x_i z}. \end{aligned} \quad \square$$

We may now use Proposition 3.8 to prove Theorem 3.2.

Proof of Theorem 3.2. First, suppose that $a \geq N$ is a positive integer. By (3.7) and Proposition 3.8, we have that

$$\frac{\mathcal{B}(\mu_{a,b}, \lambda)}{\mathcal{B}(\rho, \lambda)} = (-1)^{N-b-1} \prod_{i \neq b} \frac{b-i}{a-i} s_{a-N|N-1-b}(e^\lambda) = J(a, b) I(a, b),$$

with $I(a, b)$ and $J(a, b)$ defined by

$$\begin{aligned} I(a, b) &:= \oint_{\{0\}} \oint_{\{0\}} \frac{dz dw}{2\pi i 2\pi i} z^{b-a-1} w^{-N-1} \frac{1-w^{N-b}}{1-w} \prod_{i=1}^N \frac{w - e^{\lambda_i} z}{1 - e^{\lambda_i} z} \\ J(a, b) &:= \prod_{i \neq b} \frac{b-i}{a-i}, \end{aligned}$$

where both contours are small circles around 0. Note that the integrand has z -poles at 0 and $e^{-\lambda_i}$ and w -poles at 0 and ∞ . We deform the z -contour through infinity, so that it now encloses only $e^{-\lambda_i}$ instead

of 0 and the sign changes. Simultaneously we deform the w -contour to be a very large circle around the origin and reverse its direction so it now encloses the pole at ∞ rather than 0. Split now the double contour integral into two according to $1 - w^{N-b} = (1) + (-w^{N-b})$. On the new deformed w -contour, the first part vanishes, as there is no longer a pole at ∞ . Together, these manipulations yield

$$I(a, b) = - \oint_{\{e^{-\lambda_i}\}} \frac{dz}{2\pi i} \oint_{\{\infty\}} \frac{dw}{2\pi i} z^{-a+b-1} w^{-b-1} \frac{1}{1-w} \prod_{i=1}^N \frac{w - e^{\lambda_i} z}{1 - e^{\lambda_i} z}.$$

Let us make the change of variables

$$\tilde{z} = \frac{1}{z}, \quad \tilde{w} = \frac{w}{z}; \quad z = \frac{1}{\tilde{z}}, \quad w = \frac{\tilde{w}}{\tilde{z}}.$$

We get for positive integer $a \geq N$ that

$$(3.9) \quad I(a, b) = \oint_{\{e^{\lambda_i}\}} \frac{d\tilde{z}}{2\pi i (\tilde{z})^2} \oint_{\{\infty\}} \frac{d\tilde{w}}{2\pi i (\tilde{z})} \tilde{z}^{a-b+1+b+1} \tilde{w}^{-b-1} \frac{1}{1 - \frac{\tilde{w}}{\tilde{z}}} \prod_{i=1}^N \frac{\frac{\tilde{w}}{\tilde{z}} - e^{\lambda_i}/\tilde{z}}{1 - e^{\lambda_i}/\tilde{z}} = I'(a, b)$$

for

$$I'(a, b) = \oint_{\{e^{\lambda_i}\}} \frac{d\tilde{z}}{2\pi i} \oint_{\{0\}} \frac{d\tilde{w}}{2\pi i} \cdot \frac{\tilde{z}^a \tilde{w}^{-b-1}}{\tilde{z} - \tilde{w}} \cdot \prod_{i=1}^N \frac{\tilde{w} - e^{\lambda_i}}{\tilde{z} - e^{\lambda_i}},$$

where the $\tilde{z}$ -contour encloses the poles at e^{λ_i} and avoids the negative real axis, the $\tilde{w}$ -contour is a large circle enclosing 0 and the $\tilde{z}$ -contour, and we use one of the negative signs to switch the orientation of the $\tilde{w}$ -contour. In addition, we notice for positive integer $a \geq N$ that

$$J(a, b) = \prod_{i \neq b} \frac{b-i}{a-i} = J'(a, b) := (-1)^{N-b+1} \frac{\Gamma(N-b)\Gamma(b+1)\Gamma(a-N+1)(a-b)}{\Gamma(a+1)}.$$

Therefore, denoting the right hand side of the desired by $B(a, b) := J'(a, b)I'(a, b)$, we find that

$$(3.10) \quad \frac{\mathcal{B}(\mu_{a,b}, \lambda)}{\mathcal{B}(\rho, \lambda)} = B(a, b) \quad \text{for } a \in \{N, N+1, \dots\}.$$

We now complete the proof by checking the conditions of Carlson's theorem.

As a function of a , the expression $\frac{\mathcal{B}(\mu_{a,b}, \lambda)}{\mathcal{B}(\rho, \lambda)}$ is evidently analytic, of exponential type, and satisfies

$$\left| \frac{\mathcal{B}(\mu_{N+iy,b}, \lambda)}{\mathcal{B}(\rho, \lambda)} \right| < C e^{|y|} \quad \text{for } y \in \mathbb{R}$$

for some $C > 0$. On the other hand, notice that $J'(a, b)$ is evidently meromorphic in a with poles at $\{N-1, \dots, 0\} \setminus \{b\}$ and that by contracting first the $\tilde{z}$ -contour and then the $\tilde{w}$ -contour that $I'(a, b)$ is an analytic function in a of exponential type with $I'(N+iy, b) < C e^{|y|}$ for some $C > 0$. Further, for $a \in \{N-1, \dots, 0\} \setminus \{b\}$, by deforming the $\tilde{z}$ contour to ∞ we find that the non-residue term vanishes and we have

$$I'(a, b) = \oint_{\{0\}} \frac{d\tilde{w}}{2\pi i} \tilde{w}^{a-b-1} = 0.$$

This shows that $I'(a, b)$ has zeroes at $a \in \{N-1, \dots, 0\} \setminus \{b\}$ and hence $B(a, b)$ is analytic. Finally, $J(a, b)$ is evidently of exponential type and satisfies $J(N+iy, b) < C e^{|y|}$ for $y \in \mathbb{R}$ and some $C > 0$, meaning that both functions satisfy the conditions of Carlson's theorem as a function of $a - N$. We conclude that they are identically equal, as desired. $\square$

Remark. The definition (3.6) of Schur functions and the principal evaluation identity [Mac95, Example I.1] imply the index-variable duality formula

$$\frac{s_\lambda(q^a)}{s_\lambda(q^\rho)} = \frac{s_{a-\rho}(q^{\lambda+\rho})}{s_{a-\rho}(q^\rho)}.$$

Using this and the method of proof of Theorem 3.2, we may also obtain the identity

$$(3.11) \quad \frac{s_\lambda(1, q, \dots, q^{b-1}, q^a, q^{b+1}, \dots, q^{N-1})}{s_\lambda(1, \dots, q^{N-1})} = q^{-(N-b-1)(N-b-2)/2} \prod_{i=1}^{N-b-1} \frac{q^{N-1} - q^{N-i-1}}{q^a - q^{N-i}} \prod_{i=1}^b \frac{q^b - q^{i-1}}{q^a - q^{i-1}} \\ \times (-1)^{N-b-1} \oint_{\{q^{\lambda_i+N-i}\}} \frac{dz}{2\pi i} \oint_{\{0\}} \frac{dw}{2\pi i} \cdot \frac{z^a w^{-b-1}}{z-w} \cdot \prod_{i=1}^N \frac{w - q^{\lambda_i+N-i}}{z - q^{\lambda_i+N-i}},$$

where the z -contour encloses q^{λ_i+N-i} and avoids the negative real axis and the w -contour encloses the z -contour and 0. This identity and its generalizations to other root systems are of key importance in [CG18].

3.4. Single variable asymptotics. The proof of Theorem 3.4 will proceed through the single variable case, which gives asymptotics for the quantity

$$\tilde{B}_N(\tilde{a}, \tilde{b}) := \frac{\mathcal{B}(\mu_{\tilde{a}_N, \tilde{b}_N}, \lambda^N)}{\mathcal{B}(\rho, \lambda^N)}.$$

In this case, we will apply the method of steepest descent to derive the following specialization of Theorem 3.4 to a single variable.

Theorem 3.9. If λ^N satisfies Assumption 3.3, for $\tilde{b} \in \{0, \frac{1}{N}, \dots, \frac{N-1}{N}\}$, as a function of $\tilde{a}$, the quantity $\tilde{B}_N(\tilde{a}, \tilde{b})$ admits a holomorphic extension to an open neighborhood U of $[0, 1]$ such that uniformly on compact subsets of U and uniformly in $\tilde{b}$ we have

$$\tilde{B}_N(\tilde{a}, \tilde{b}) = \frac{1}{M_{\tilde{\rho}_N}^{-1}(\tilde{a}-1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}-1)} \frac{\tilde{a} - \tilde{b}}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a}-1))M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b}-1))}} \frac{\sqrt{S_{\tilde{\rho}_N}(\tilde{b}-1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{b})}}{\sqrt{S_{\tilde{\rho}_N}(\tilde{a}-1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{a})}} [1+o(1)].$$

Remark. By Lemma 3.5, the expression

$$\frac{1}{M_{\tilde{\rho}_N}^{-1}(\tilde{a}-1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}-1)} \frac{\tilde{a} - \tilde{b}}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a}-1))M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b}-1))}} \frac{\sqrt{S_{\tilde{\rho}_N}(\tilde{b}-1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{b})}}{\sqrt{S_{\tilde{\rho}_N}(\tilde{a}-1)}e^{N\tilde{\Psi}_{\rho_N}(\tilde{a})}}$$

in Theorem 3.9 is analytic in $\tilde{a}$ on a complex neighborhood of $[0, 1]$.

To prove Theorem 3.9, we need a few asymptotic preliminaries.

Lemma 3.10 ([BV92, Proposition 3.1 and 3.3]). There is a neighborhood $U \supset [0, 1]$ such that $M_{\tilde{\rho}_N}^{-1}(z-1)$ is well-defined, bijective, and meromorphic on U with unique pole at 1 and zero at 0 and $S_{\tilde{\rho}_N}(z-1)$ is holomorphic with no zeros on U . Further, the functions $M_{\tilde{\rho}_N}^{-1}(z-1)$ converge uniformly on compact subsets of $U \setminus \{1\}$ to $M_{\tilde{\rho}}^{-1}(z-1)$, and the functions $S_{\tilde{\rho}_N}(z-1)$ converge uniformly on compact subsets of U to $S_{\tilde{\rho}}(z-1)$.

Lemma 3.11. We have that $S_{\tilde{\rho}_N}(z) > 0$ for $z \in [-1, 0]$ and that

$$S_{\tilde{\rho}_N}(0) = \left[\int e^s d\rho_N(s) \right]^{-1} \quad \text{and} \quad S_{\tilde{\rho}_N}(-1) = \int e^{-s} d\rho_N(s).$$

Proof. The first statement follows by [BV92, Proposition 3.1(3)] and the latter two are straightforward by series expansion in the definition of $S_{\tilde{\rho}_N}(u)$. $\square$

Proof of Theorem 3.9. The argument is based on the steepest descent method, cf. [Erd56, Cop65]. We deform the z - and w -contours in the double contour integral representation of Theorem 3.2 so that they pass through the critical points the z and w parts of the integrand (multiplied by $(z-w)$). The leading contribution to the asymptotics is then given by integrals over small neighborhoods of these critical points, where we can consider Taylor expansions of the integrand. We now present the technical details. Define the function

$$h_x(y) := xy - \int \log(e^y - e^s) d\rho_N(s)$$

for which we have that

$$\begin{aligned} h'_x(y) &= x - 1 - M_{\tilde{\rho}_N}(e^{-y}) \\ h''_x(y) &= e^{-y} M'_{\tilde{\rho}_N}(e^{-y}) \\ h'''_x(y) &= -e^{-y} M''_{\tilde{\rho}_N}(e^{-y}) - e^{-2y} M'''_{\tilde{\rho}_N}(e^{-y}). \end{aligned}$$

In particular, for $x \in U \setminus \{0, 1\}$, by Lemma 3.10, $h_x(y)$ has critical points at $-\log M_{\tilde{\rho}_N}^{-1}(x-1) + 2\pi i\mathbb{Z}$. We define

$$y_x := -\log M_{\tilde{\rho}_N}^{-1}(x-1)$$

to be the unique critical point of this form with $\Im(y_x) \in [0, 2\pi)$. Notice in particular that

$$(3.12) \quad h''_x(y_x + t) = \int \frac{e^{s+t} M_{\tilde{\rho}_N}^{-1}(x-1)}{(1 - e^{s+t} M_{\tilde{\rho}_N}^{-1}(x-1))^2} d\rho_N(s)$$

$$(3.13) \quad h'''_x(y_x + t) = \int \left[\frac{e^{s+t} M_{\tilde{\rho}_N}^{-1}(x-1)(1 + e^{s+t} M_{\tilde{\rho}_N}^{-1}(x-1))}{(1 - e^{s+t} M_{\tilde{\rho}_N}^{-1}(x-1))^3} \right] d\rho_N(s).$$

For $1/2 > \varepsilon > 0$, define the counterclockwise contour

$$\begin{aligned} \Upsilon_\varepsilon := \{x - i\varepsilon \mid x \in [0, 1]\} \cup \{1 + \varepsilon \exp(2\pi i x) \mid x \in [-1/4, 1/4]\} \cup \\ \{x + i\varepsilon \mid x \in [0, 1]\} \cup \{\varepsilon \exp(2\pi i x) \mid x \in [1/4, 3/4]\} \end{aligned}$$

shown in Figure 3.1 which encloses $[0, 1]$ and maintains a distance ε from it.

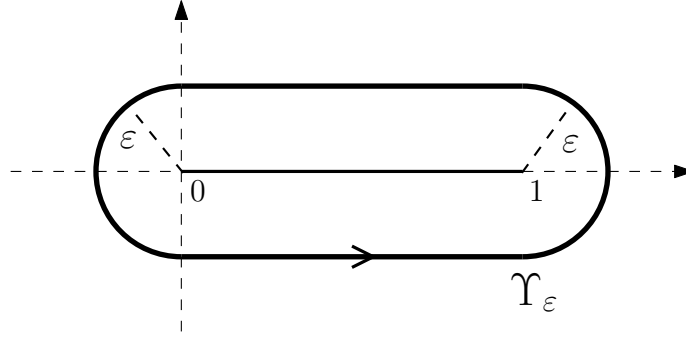


FIGURE 3.1. The contour Υ_ε

Plugging in the value of y_x , we obtain the expressions

$$h''_x(y_x) = \frac{M_{\tilde{\rho}_N}^{-1}(x-1)}{\partial_x M_{\tilde{\rho}_N}^{-1}(x-1)} \quad \text{and} \quad h'''_x(y_x) = [\partial_x^2 M_{\tilde{\rho}_N}^{-1}(x-1)][\partial_x M_{\tilde{\rho}_N}^{-1}(x-1)]^3 [M_{\tilde{\rho}_N}^{-1}(x-1)]^2 - \frac{M_{\tilde{\rho}_N}^{-1}(x-1)}{\partial_x M_{\tilde{\rho}_N}^{-1}(x-1)}$$

from which Lemma 3.10 implies that $h''_x(y_x)$ and $h'''_x(y_x)$ are meromorphic as functions of x . Furthermore, for $x \in [0, 1]$, by the expression (3.12) and the fact that $M_{\tilde{\rho}_N}^{-1}(x-1)$ is non-positive, we see that $h''_x(y_x)$ is differentiable in x and $h''_x(y_x) \leq 0$ with zeroes only at $x \in \{0, 1\}$. In addition, by the expression (3.13) we see that $h'''_x(y_x)$ is differentiable and bounded on $x \in [0, 1]$, with zeroes at $x \in \{0, 1\}$. By Lemma 3.10, all of these properties hold uniformly in N for large enough N . We conclude that there exists $C_1, C'_1, C''_1 > 0$, $\varepsilon > 0$, and $\tau > 0$ so that for large enough N we have uniformly in N and $\tilde{b}$ by the fact that $|h''_{\tilde{a}}(y_{\tilde{a}})|$ is bounded for $\tilde{a} \in [0, 1]$ that

$$(3.14) \quad |h''_{\tilde{a}}(y_{\tilde{a}})| \in [C_1^{-1}, C_1] \quad \text{for all } \tilde{a} \in \Upsilon_\varepsilon$$

and by the boundedness of $|h'''_{\tilde{a}}(y_{\tilde{a}})|$ for $\tilde{a} \in [0, 1]$ that

$$(3.15) \quad |h'''_{\tilde{a}}(y_{\tilde{a}})| \leq C_1 \quad \text{for all } \tilde{a} \in \Upsilon_\varepsilon$$

and by the boundedness of $|h''_{\tilde{b}}(y_{\tilde{b}})|$ for $\tilde{b} \in [0, 1]$ that

$$(3.16) \quad h''_{\tilde{b}}(y_{\tilde{b}}) \in [-C_1, -C_1^{-1}] \quad \text{for all } \tilde{b} \in [\tau, 1 - \tau]$$

and by the fact that $h_b''(y_b)$ is non-positive on $[0, 1]$ and has zeroes and is differentiable at $\tilde{b} \in \{0, 1\}$ that

$$(3.17) \quad h_b''(y_b) \leq -C_1' \tilde{b} \text{ for } \tilde{b} \in [0, \tau] \quad h_b''(y_b) \leq -C_1'(1 - \tilde{b}) \text{ for } \tilde{b} \in [1 - \tau, 1]$$

and by the boundedness of $h_b'''(y_b)$ for $\tilde{b} \in [0, 1]$ that

$$(3.18) \quad |h_b'''(y_b)| \leq C_1 \quad \text{for } \tilde{b} \in [\tau, 1 - \tau]$$

and by (3.13) and Lemma 3.10 that uniformly in $|t| < \tau$, we have

$$(3.19) \quad |h_b'''(y_b + t)| \leq C_1' \tilde{b} \text{ for } \tilde{b} \in [0, \tau] \quad |h_b'''(y_b + t)| \leq C_1'(1 - \tilde{b}) \text{ for } \tilde{b} \in [1 - \tau, 1].$$

From now on, we fix this $\varepsilon > 0$. Define the domain V_ε to be the interior of the region enclosed by Υ_ε .

Notice that both $\tilde{B}_N(\tilde{a}, \tilde{b})$ and the claimed asymptotic expression are analytic in $\tilde{a}$ for each N by Lemma 3.5. Therefore, if we can show uniform convergence for $\tilde{a} \in \Upsilon_\varepsilon$, then this will show that their differences form a sequence of holomorphic functions on a neighborhood of $\Upsilon_\varepsilon \cup V_\varepsilon$ which converge to 0 on Υ_ε . By the Cauchy integral formula, this would imply that the difference also converges to 0 on V_ε . Thus, to prove the desired statement for $U = V_\varepsilon$, it suffices for us to prove it for $\tilde{a} \in \Upsilon_\varepsilon$. For the rest of the proof, we will therefore assume that $\tilde{a} \in \Upsilon_\varepsilon$.

We now define

$$I(\tilde{a}, \tilde{b}) := \oint_{\{e^{\lambda_i^N}\}} \frac{dz}{2\pi i} \oint_{\{0\}} \frac{dw}{2\pi i} \frac{z^{\tilde{a}N} w^{-\tilde{b}N-1}}{z - w} \prod_{i=1}^N \frac{w - e^{\lambda_i^N}}{z - e^{\lambda_i^N}}$$

so that

$$\tilde{B}_N(\tilde{a}, \tilde{b}) = (-1)^{(1-\tilde{b})N+1} \frac{\Gamma(N - \tilde{b}N) \Gamma(\tilde{b}N + 1) \Gamma(\tilde{a}N - N + 1) (\tilde{a} - \tilde{b})^N}{\Gamma(\tilde{a}N + 1)} I(\tilde{a}, \tilde{b})$$

by Theorem 3.2. Choose $M > 1$ so that $\max\{e^s, e^{-s}\} < M$ for $s \in I$, where I is the interval from Assumption 3.3. Given this M , choose k_N so that as $N \rightarrow \infty$ we have

$$(3.20) \quad k_N = \omega(1) \quad \text{and} \quad M^{k_N} \frac{k_N^2}{N} = o(N^{1/2}).$$

We divide the analysis into several cases depending on the value of $\tilde{b}$, where we wish to obtain convergence uniform in $\tilde{b}$ in each case.

Case 1: Suppose that $k_N/N < \tilde{b} < 1 - k_N/N$. Make the change of variables $w = e^{\tilde{w}}$ and $z = e^{\tilde{z}}$ so that

$$I(\tilde{a}, \tilde{b}) = \oint_{\Gamma'_a} \frac{d\tilde{z}}{2\pi i} \oint_{\Sigma'_b} \frac{d\tilde{w}}{2\pi i} \frac{1}{1 - e^{\tilde{w} - \tilde{z}}} \exp\left(N(h_{\tilde{a}}(\tilde{z}) - h_{\tilde{b}}(\tilde{w}))\right),$$

where the $\tilde{z}$ -contour Γ'_a is a positively oriented contour around $\{\lambda_i^N\}$ and the $\tilde{w}$ -contour Σ'_b connects y_b to $y_b + 2\pi i$ while staying to the right of Γ'_a .

Case 1, Step 1. We deform Γ'_a and Σ'_b to steepest descent contours. For Σ'_b , notice that $h_b(y)$ has a unique critical point at y_b . In addition, because $\tilde{b} \in (0, 1)$, we see that $y_b \in (-\infty, 0) + \pi i$ and that

$$\Re h_b(y) = \tilde{b} \cdot \Re[y] - \int \log |e^y - e^s| d\rho_N(s),$$

which implies that on the vertical line segment passing through y_b , $\Re h_b(y)$ is minimized at $y \in y_b + 2\pi i\mathbb{Z}$. Putting these facts together, we deform Σ'_b to the line segment Σ_b between y_b and $y_b + 2\pi i$ so that the resulting contour Σ_b satisfies the following properties uniformly in N :

- Σ_b lies within $\{\Re h_b(y) > \Re h_b(y_b)\}$ away from its endpoints, where we note that $\Re h_b(y)$ may be infinite on Σ_b ;
- for points on Σ_b , in constant-size neighborhood of y_b and $y_b + 2\pi i$, the value of $\Re[h_b(y) - h_b(y_b)]$ is smaller on this neighborhood than its complement;
- Σ_b has bounded length.

For Γ'_a , notice that $h_a(y)$ has a unique critical point at y_a and that $\Re h_a(y)$ is harmonic away from $I + 2\pi i\mathbb{Z}$. Therefore, at least one of the components of $\{\Re h_a(y) > \Re h_a(y_a)\}$ has compact closure and contains either

$\text{supp } d\rho_N$ or $\text{supp } d\rho_N + 2\pi\mathbf{i}$. In addition, for $\eta > 0$ to be determined later, we see by the weak convergence of $d\rho_N$ to $d\rho$ that $\Re h_{\tilde{a}}(y)$ converges uniformly on compact subsets of

$$D_\eta := \mathbb{C} \setminus \{z \mid |z - w| > \eta \text{ for } w \in I + 2\pi\mathbf{i}\mathbb{Z}\}$$

to the function $\Re h_{\tilde{a}}^\infty(y)$ for

$$h_{\tilde{a}}^\infty(y) := \tilde{a}y - \int \log(e^y - e^s) d\rho(s).$$

As a consequence, the intersection of the level set $\{\Re h_{\tilde{a}}(y) = \Re h_{\tilde{a}}(y_{\tilde{a}})\}$ with D_η converges to a fixed contour. Because $y_{\tilde{a}} \in D_\eta$ for small enough η , for small enough $\eta > 0$, there exists some small $\nu > 0$ so that the line through $y_{\tilde{a}}$ in the direction of $[-h_{\tilde{a}}''(y_{\tilde{a}})]^{-1/2}$ intersects $\{\Re h_{\tilde{a}}(y) = \Re h_{\tilde{a}}(y_{\tilde{a}}) - \nu\}$ at least a fixed distance away from $y_{\tilde{a}}$ for large enough N . Notice that the union of $\{\Re h_{\tilde{a}}(y) = \Re h_{\tilde{a}}(y_{\tilde{a}}) - \nu\}$ and the line through $y_{\tilde{a}}$ in the direction of $[-h_{\tilde{a}}''(y_{\tilde{a}})]^{-1/2}$ contains a loop which encloses either $\text{supp } d\rho_N$ or $\text{supp } d\rho_N + 2\pi\mathbf{i}$. We deform $\Gamma_{\tilde{a}}'$ to the closed contour $\Gamma_{\tilde{a}}$ defined as the pieces of the contour forming this loop.

Notice that $\Gamma_{\tilde{a}} \cap D_\eta$ converges to a fixed contour, while $\Gamma_{\tilde{a}} \cap (\mathbb{C} \setminus D_\eta)$ lies within a single connected component of $\mathbb{C} \setminus D_\eta$, hence has bounded length. We conclude that $\Gamma_{\tilde{a}}$ satisfies the following properties uniformly in N :

- in a constant size neighborhood of $y_{\tilde{a}}$, $\Gamma_{\tilde{a}}$ is a line segment in the direction of $[-h_{\tilde{a}}''(y_{\tilde{a}})]^{-1/2}$, and the value of $\Re[h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{a}}(y)]$ is smaller on this neighborhood than its complement;
- $\Gamma_{\tilde{a}}$ is a closed loop passing through $y_{\tilde{a}}$ which encloses $\text{supp } d\rho_N$ or $\text{supp } d\rho_N + 2\pi\mathbf{i}$;
- $\Gamma_{\tilde{a}}$ lies within $\{\Re h_{\tilde{a}}(y) < \Re h_{\tilde{a}}(y_{\tilde{a}})\}$ outside of $y_{\tilde{a}}$ and D_η ;
- $\Gamma_{\tilde{a}}$ has bounded length.

We illustrate the two contours $\Sigma_{\tilde{b}}$ and $\Gamma_{\tilde{a}}$ in Figure 3.2. After these two constructions, we see that

- $\Re h_{\tilde{b}}(y)$ achieves a global minimum on $\Sigma_{\tilde{b}}$ at $y_{\tilde{b}}$ and $y_{\tilde{b}} + 2\pi\mathbf{i}$;
- $\Re h_{\tilde{a}}(y)$ achieves a global maximum on $\Gamma_{\tilde{a}}$ at $y_{\tilde{a}}$.

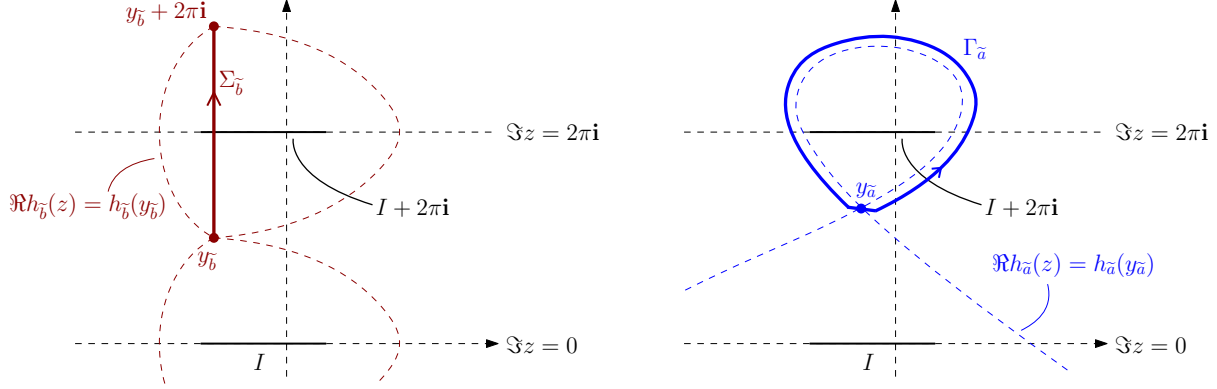


FIGURE 3.2. Schematic configuration of the contours $\Sigma_{\tilde{b}}$ and $\Gamma_{\tilde{a}}$

In the course of the deformations described above, we may pick up some residues. Denoting these residues by $R(\tilde{a}, \tilde{b})$, we find that

$$I(\tilde{a}, \tilde{b}) = I'(\tilde{a}, \tilde{b}) + R(\tilde{a}, \tilde{b})$$

for

$$I'(\tilde{a}, \tilde{b}) := \oint_{\Gamma_{\tilde{a}}} \frac{d\tilde{z}}{2\pi\mathbf{i}} \oint_{\Sigma_{\tilde{b}}} \frac{d\tilde{w}}{2\pi\mathbf{i}} \frac{1}{1 - e^{\tilde{w} - \tilde{z}}} \exp\left(N(h_{\tilde{a}}(\tilde{z}) - h_{\tilde{b}}(\tilde{w}))\right)$$

$$R(\tilde{a}, \tilde{b}) := \oint_{\Xi} \frac{d\tilde{w}}{2\pi\mathbf{i}} \exp\left(N(\tilde{a} - \tilde{b})\tilde{w}\right),$$

where Ξ is a closed contour which is either empty or has endpoints on $\Gamma_{\tilde{a}} \cap \Sigma_{\tilde{b}}$, and we adopt the convention that $\exp(Nh_{\tilde{b}}(\tilde{w})) = 0$ if $\Re h_{\tilde{b}}(\tilde{w}) = \infty$. Note that the contours $\Gamma_{\tilde{a}}, \Sigma_{\tilde{b}}, \Xi$ may depend on N , but in what

follows, we will use only the properties specified above, which hold for all large enough N . Define $\delta_{\tilde{a}} = \delta_{\tilde{a}}(N) > 0$ and $\delta_{\tilde{b}} = \delta_{\tilde{b}}(N) > 0$ by

$$(3.21) \quad \delta_{\tilde{a}} = N^{-2/5} \quad \delta_{\tilde{b}} = N^{-2/5} |h''_{\tilde{b}}(y_{\tilde{b}})|^{-2/5}$$

so that $\delta_{\tilde{b}} = o(1)$ by (3.16), (3.17), and the fact that $\tilde{b} \in [k_N/N, 1 - k_N/N]$. For these choices, define the decompositions $\Gamma_{\tilde{a}} = \Gamma_{\tilde{a}}^1 \sqcup \Gamma_{\tilde{a}}^2$ and $\Sigma_{\tilde{b}} = \Sigma_{\tilde{b}}^1 \sqcup \Sigma_{\tilde{b}}^2$ for

$$\Gamma_{\tilde{a}}^1 = B(y_{\tilde{a}}, \delta_{\tilde{a}}) \cap \Gamma_{\tilde{a}}, \quad \Gamma_{\tilde{a}}^2 = \Gamma_{\tilde{a}} \setminus \Gamma_{\tilde{a}}^1, \quad \Sigma_{\tilde{b}}^1 = [B(y_{\tilde{b}}, \delta_{\tilde{b}}) \cup B(y_{\tilde{b}} + 2\pi\mathbf{i}, \delta_{\tilde{b}})] \cap \Sigma_{\tilde{b}}, \quad \Sigma_{\tilde{b}}^2 = \Sigma_{\tilde{b}} \setminus \Sigma_{\tilde{b}}^1.$$

Since $\delta_{\tilde{a}}^2 = o(1)$ and $\delta_{\tilde{b}}^2 |h''_{\tilde{b}}(y_{\tilde{b}})| = o(1)$, for large enough N , we have uniformly in $\tilde{a}, \tilde{b}$ that

$$(3.22) \quad \Re[h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{a}}(y)] > \frac{1}{4} |h''_{\tilde{a}}(y_{\tilde{a}})| \delta_{\tilde{a}}^2 \quad \text{for } y \in \Gamma_{\tilde{a}}^2$$

$$(3.23) \quad \Re[h_{\tilde{b}}(y) - h_{\tilde{b}}(y_{\tilde{b}})] > \frac{1}{4} |h''_{\tilde{b}}(y_{\tilde{b}})| \delta_{\tilde{b}}^2 \quad \text{for } y \in \Sigma_{\tilde{b}}^2.$$

We now analyze $I'(\tilde{a}, \tilde{b})$ and $R(\tilde{a}, \tilde{b})$ separately.

Case 1, Step 2. We claim that

$$(3.24) \quad I'(\tilde{a}, \tilde{b}) = \frac{-\mathbf{i}}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \frac{1}{2\pi N \sqrt{h''_{\tilde{a}}(y_{\tilde{a}}) h''_{\tilde{b}}(y_{\tilde{b}})}} e^{N[h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{b}}(y_{\tilde{b}})]} (1 + o(1)),$$

where we take the standard branch of the square root. Consider the decomposition

$$I'(\tilde{a}, \tilde{b}) = I_1(\tilde{a}, \tilde{b}) + I_2(\tilde{a}, \tilde{b}) + I_3(\tilde{a}, \tilde{b}) + I_4(\tilde{a}, \tilde{b})$$

for

$$\begin{aligned} I_1(\tilde{a}, \tilde{b}) &:= \frac{1}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \left[\oint_{\Gamma_{\tilde{a}}^1} \exp(Nh_{\tilde{a}}(\tilde{z})) \frac{d\tilde{z}}{2\pi\mathbf{i}} \right] \left[\oint_{\Sigma_{\tilde{b}}^1} \exp(-N(h_{\tilde{b}}(\tilde{w})) \frac{d\tilde{w}}{2\pi\mathbf{i}} \right] \\ I_2(\tilde{a}, \tilde{b}) &:= \oint_{\Gamma_{\tilde{a}}^2} \frac{d\tilde{z}}{2\pi\mathbf{i}} \oint_{\Sigma_{\tilde{b}}^1} \frac{d\tilde{w}}{2\pi\mathbf{i}} \frac{1}{1 - e^{\tilde{w} - \tilde{z}}} \exp(N(h_{\tilde{a}}(\tilde{z}) - h_{\tilde{b}}(\tilde{w}))) \\ I_3(\tilde{a}, \tilde{b}) &:= \oint_{\Gamma_{\tilde{a}}^1} \frac{d\tilde{z}}{2\pi\mathbf{i}} \oint_{\Sigma_{\tilde{b}}^2} \frac{d\tilde{w}}{2\pi\mathbf{i}} \frac{1}{1 - e^{\tilde{w} - \tilde{z}}} \exp(N(h_{\tilde{a}}(\tilde{z}) - h_{\tilde{b}}(\tilde{w}))) \\ I_4(\tilde{a}, \tilde{b}) &:= \oint_{\Gamma_{\tilde{a}}^1} \frac{d\tilde{z}}{2\pi\mathbf{i}} \oint_{\Sigma_{\tilde{b}}^1} \frac{d\tilde{w}}{2\pi\mathbf{i}} \left[\frac{1}{1 - e^{\tilde{w} - \tilde{z}}} - \frac{1}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \right] \exp(N(h_{\tilde{a}}(\tilde{z}) - h_{\tilde{b}}(\tilde{w}))) \end{aligned}$$

where we note that $y_{\tilde{a}} = -\log M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1)$ and $y_{\tilde{b}} = -\log M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)$ are not equal because the domains Υ_ε and $\{0, 1/N, \dots, (N-1)/N\}$ of $\tilde{a}$ and $\tilde{b}$ do not intersect and $M_{\tilde{\rho}_N}^{-1}(u - 1)$ is bijective on U by Lemma 3.10, meaning that $1 - e^{y_{\tilde{b}} - y_{\tilde{a}}} \neq 0$. We first analyze $I_1(\tilde{a}, \tilde{b})$. For $x \in \{\tilde{a}, \tilde{b}\}$ and $|y - y_x| < \delta_x$, we have the Taylor expansion

$$h_x(y) = h_x(y_x) + \frac{1}{2} h''_x(y_x)(y - y_x)^2 + \frac{1}{6} h'''_x(\xi_x(y))(y - y_x)^3,$$

where $\xi_x(y) \in B(y_x, \delta_x)$. Consider the resulting decomposition

$$(3.25) \quad J(\tilde{a}) := \oint_{\Gamma_{\tilde{a}}^1} \exp(Nh_{\tilde{a}}(\tilde{z})) \frac{d\tilde{z}}{2\pi\mathbf{i}} = J_1(\tilde{a}) - J_2(\tilde{a}) + J_3(\tilde{a})$$

for

$$\begin{aligned} J_1(\tilde{a}) &:= e^{Nh_{\tilde{a}}(y_{\tilde{a}})} \oint_{\ell_{\tilde{a}}} e^{\frac{N}{2} h''_{\tilde{a}}(y_{\tilde{a}})(\tilde{z} - y_{\tilde{a}})^2} \frac{d\tilde{z}}{2\pi\mathbf{i}} \\ J_2(\tilde{a}) &:= e^{Nh_{\tilde{a}}(y_{\tilde{a}})} \oint_{\ell_{\tilde{a}} - \Gamma_{\tilde{a}}^1} e^{\frac{N}{2} h''_{\tilde{a}}(y_{\tilde{a}})(\tilde{z} - y_{\tilde{a}})^2} \frac{d\tilde{z}}{2\pi\mathbf{i}} \\ J_3(\tilde{a}) &:= e^{Nh_{\tilde{a}}(y_{\tilde{a}})} \oint_{\Gamma_{\tilde{a}}^1} e^{\frac{N}{2} h''_{\tilde{a}}(y_{\tilde{a}})(\tilde{z} - y_{\tilde{a}})^2} \left(e^{\frac{N}{6} h'''_{\tilde{a}}(\xi_{\tilde{a}}(\tilde{z}))(\tilde{z} - y_{\tilde{a}})^3} - 1 \right) \frac{d\tilde{z}}{2\pi\mathbf{i}}, \end{aligned}$$

where $\ell_{\tilde{a}}$ is the line passing through $y_{\tilde{a}}$ in the direction of $[-h''_{\tilde{a}}(y_{\tilde{a}})]^{-1/2}$. The main contributing factor will be

$$J_1(\tilde{a}) = e^{Nh_{\tilde{a}}(y_{\tilde{a}})} [-h''_{\tilde{a}}(y_{\tilde{a}})]^{-1/2} \int_{-\infty}^{\infty} e^{-\frac{N}{2}t^2} \frac{dt}{2\pi\mathbf{i}} = \frac{-\mathbf{i}}{\sqrt{2\pi N h''_{\tilde{a}}(y_{\tilde{a}})}} e^{Nh_{\tilde{a}}(y_{\tilde{a}})}.$$

In this expression, notice that

$$h''_{\tilde{a}}(y_{\tilde{a}}) = M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1) M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1)) = \frac{\tilde{a} - 1}{\tilde{a}} S_{\tilde{\rho}_N}(\tilde{a} - 1) M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1)),$$

where the loop each term in the final expression traces out for $\tilde{a} \in \Upsilon_\varepsilon$ does not enclose 0, meaning that we may choose a branch of the logarithm such that $\sqrt{h''_{\tilde{a}}(y_{\tilde{a}})}$ is well-defined for all $\tilde{a} \in \Upsilon_\varepsilon$. We now bound each other term relative to $|J_1(\tilde{a})|$. Notice that for large enough N , we have by a Gaussian tail bound that

$$\begin{aligned} |J_2(\tilde{a})| &\leq |e^{Nh_{\tilde{a}}(y_{\tilde{a}})}| |h''_{\tilde{a}}(y_{\tilde{a}})|^{-1/2} \int_{(-\infty, -\delta_{\tilde{a}}|h''_{\tilde{a}}(y_{\tilde{a}})|^{1/2}] \cup [\delta_{\tilde{a}}|h''_{\tilde{a}}(y_{\tilde{a}})|^{1/2}, \infty)} |e^{-\frac{N}{2}t^2}| \frac{dt}{2\pi} \\ &\leq 2|e^{Nh_{\tilde{a}}(y_{\tilde{a}})}| |h''_{\tilde{a}}(y_{\tilde{a}})|^{-1/2} \cdot e^{-\frac{N}{2}\delta_{\tilde{a}}^2|h''_{\tilde{a}}(y_{\tilde{a}})|} \\ &\leq e^{-C_2 N^{1/5}} |J_1(\tilde{a})| \end{aligned}$$

for some $C_2 > 0$, where the last inequality follows from (3.14) and (3.21). Similarly, we have for large enough N that

$$\begin{aligned} |J_3(\tilde{a})| &\leq |e^{Nh_{\tilde{a}}(y_{\tilde{a}})}| |h''_{\tilde{a}}(y_{\tilde{a}})|^{-1/2} \int_{-\delta_{\tilde{a}}|h''_{\tilde{a}}(y_{\tilde{a}})|^{1/2}}^{\delta_{\tilde{a}}|h''_{\tilde{a}}(y_{\tilde{a}})|^{1/2}} e^{-\frac{N}{2}t^2} \left| e^{\frac{N}{6}h'''_{\tilde{a}}(\xi_{\tilde{a}}(y_{\tilde{a}} + [-h''_{\tilde{a}}(y_{\tilde{a}})]^{-1/2}t))} [-h''_{\tilde{a}}(y_{\tilde{a}})]^{-3/2}t^3 - 1 \right| \frac{dt}{2\pi} \\ &\leq |e^{Nh_{\tilde{a}}(y_{\tilde{a}})}| |h''_{\tilde{a}}(y_{\tilde{a}})|^{-1/2} N^{-1/2} \int_{-\delta_{\tilde{a}}|h''_{\tilde{a}}(y_{\tilde{a}})|^{1/2}\sqrt{N}}^{\delta_{\tilde{a}}|h''_{\tilde{a}}(y_{\tilde{a}})|^{1/2}\sqrt{N}} e^{-\frac{1}{2}s^2} \left| e^{\frac{1}{6}h'''_{\tilde{a}}(\xi_{\tilde{a}}(y_{\tilde{a}} + [-h''_{\tilde{a}}(y_{\tilde{a}})]^{-1/2}t))} [-h''_{\tilde{a}}(y_{\tilde{a}})]^{-3/2}s^3 N^{-1/2} - 1 \right| \frac{ds}{2\pi} \\ &\leq C_3 \delta_{\tilde{a}}^3 N^{1/2} |J_1(\tilde{a})| \\ &= C_3 N^{-7/10} |J_1(\tilde{a})| \end{aligned}$$

for some $C_3 > 0$, where we use (3.15) and (3.21) in the last two inequalities. We conclude that

$$J(\tilde{a}) = \frac{-\mathbf{i}}{\sqrt{2\pi N h''_{\tilde{a}}(y_{\tilde{a}})}} e^{Nh_{\tilde{a}}(y_{\tilde{a}})} [1 + o(1)].$$

Consider now

$$(3.26) \quad K(\tilde{b}) := \oint_{\Sigma_b^\perp} \exp(-N(h_{\tilde{b}}(\tilde{w}))) \frac{d\tilde{w}}{2\pi\mathbf{i}} = K_1(\tilde{b}) - K_2(\tilde{b}) + K_3(\tilde{b})$$

for

$$\begin{aligned} K_1(\tilde{b}) &:= e^{-Nh_{\tilde{b}}(y_{\tilde{b}})} \left[\oint_{\ell_b^+} e^{-\frac{N}{2}h''_{\tilde{b}}(y_{\tilde{b}})(\tilde{w}-y_{\tilde{b}})^2} \frac{d\tilde{w}}{2\pi\mathbf{i}} + \oint_{\ell_b^-} e^{-\frac{N}{2}h''_{\tilde{b}}(y_{\tilde{b}})(\tilde{w}-y_{\tilde{b}}-2\pi\mathbf{i})^2} \frac{d\tilde{w}}{2\pi\mathbf{i}} \right] \\ K_2(\tilde{b}) &:= e^{-Nh_{\tilde{b}}(y_{\tilde{b}})} \left[\oint_{\ell_b^+ - B(y_{\tilde{b}}, \delta)} e^{-\frac{N}{2}h''_{\tilde{b}}(y_{\tilde{b}})(\tilde{w}-y_{\tilde{b}})^2} \frac{d\tilde{w}}{2\pi\mathbf{i}} + \oint_{\ell_b^- - B(y_{\tilde{b}}+2\pi\mathbf{i}, \delta)} e^{-\frac{N}{2}h''_{\tilde{b}}(y_{\tilde{b}})(\tilde{w}-y_{\tilde{b}}-2\pi\mathbf{i})^2} \frac{d\tilde{w}}{2\pi\mathbf{i}} \right] \\ K_3(\tilde{b}) &:= e^{-Nh_{\tilde{b}}(y_{\tilde{b}})} \left[\oint_{\Sigma_b^\perp \cap B(y_{\tilde{b}}, \delta)} e^{-\frac{N}{2}h''_{\tilde{b}}(y_{\tilde{b}})(\tilde{w}-y_{\tilde{b}})^2} \left(e^{\frac{N}{6}h'''_{\tilde{b}}(\xi_{\tilde{b}}(\tilde{w}))(\tilde{w}-y_{\tilde{b}})^3} - 1 \right) \frac{d\tilde{w}}{2\pi\mathbf{i}} \right. \\ &\quad \left. + \oint_{\Sigma_b^\perp \cap B(y_{\tilde{b}}+2\pi\mathbf{i}, \delta)} e^{-\frac{N}{2}h''_{\tilde{b}}(y_{\tilde{b}})(\tilde{w}-y_{\tilde{b}}-2\pi\mathbf{i})^2} \left(e^{\frac{N}{6}h'''_{\tilde{b}}(\xi_{\tilde{b}}(\tilde{w}))(\tilde{w}-y_{\tilde{b}}-2\pi\mathbf{i})^3} - 1 \right) \frac{d\tilde{w}}{2\pi\mathbf{i}} \right], \end{aligned}$$

where ℓ_b^+ is the upwards vertical ray through $y_{\tilde{b}}$, ℓ_b^- is the downwards vertical ray through $y_{\tilde{b}} + 2\pi\mathbf{i}$, and we note $h''_{\tilde{b}}(y + 2\pi\mathbf{i}) = h''_{\tilde{b}}(y)$ and $N[h_{\tilde{b}}(y_{\tilde{b}} + 2\pi\mathbf{i}) - h_{\tilde{b}}(y_{\tilde{b}})] = 2\pi\mathbf{i}\tilde{b} \in 2\pi\mathbf{i}\mathbb{Z}$ for $\tilde{b}N \in \mathbb{Z}$. As before, the main contributing factor is

$$K_1(\tilde{b}) = e^{-Nh_{\tilde{b}}(y_{\tilde{b}})} [h''_{\tilde{b}}(y_{\tilde{b}})]^{-1/2} \int_{-\infty}^{\infty} e^{-\frac{N}{2}t^2} \frac{dt}{2\pi\mathbf{i}} = \frac{1}{\sqrt{2\pi N h''_{\tilde{b}}(y_{\tilde{b}})}} e^{-Nh_{\tilde{b}}(y_{\tilde{b}})}.$$

For large enough N , we have by a Gaussian tail bound that

$$\begin{aligned}
|K_2(\tilde{b})| &\leq e^{-N h_{\tilde{b}}(y_{\tilde{b}})} |h_{\tilde{b}}''(y_{\tilde{b}})|^{-1/2} \int_{(-\infty, -\delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2}] \cup [\delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2}, \infty)} e^{-\frac{N}{2} t^2} \frac{dt}{2\pi} \\
&\leq 2e^{-N h_{\tilde{b}}(y_{\tilde{b}})} |h_{\tilde{b}}''(y_{\tilde{b}})|^{-1/2} e^{-\frac{N}{2} \delta_{\tilde{b}}^2 |h_{\tilde{b}}''(y_{\tilde{b}})|} \\
&= 2e^{-N h_{\tilde{b}}(y_{\tilde{b}})} |h_{\tilde{b}}''(y_{\tilde{b}})|^{-1/2} e^{-[N |h_{\tilde{b}}''(y_{\tilde{b}})|]^{1/5}/2} \\
&= |K_1(\tilde{b})| o(1),
\end{aligned}$$

where we apply (3.21) and the fact that $N|h_{\tilde{b}}''(\tilde{b})| = \omega(1)$. Similarly, for large enough N , we have

$$\begin{aligned}
|K_3(\tilde{b})| &\leq e^{-N h_{\tilde{b}}(y_{\tilde{b}})} |h_{\tilde{b}}''(y_{\tilde{b}})|^{-1/2} \left[\int_0^{\delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2}} e^{-\frac{N}{2} t^2} \left| e^{-\mathbf{i} \frac{N}{6} h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} t))} [-h_{\tilde{b}}''(y_{\tilde{b}})]^{-3/2} t^3 - 1 \right| \frac{dt}{2\pi} \right. \\
&\quad \left. + \int_{-\delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2}}^0 e^{-\frac{N}{2} t^2} \left| e^{-\mathbf{i} \frac{N}{6} h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + 2\pi\mathbf{i} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} t))} [-h_{\tilde{b}}''(y_{\tilde{b}})]^{-3/2} t^3 - 1 \right| \frac{dt}{2\pi} \right] \\
&\leq e^{-N h_{\tilde{b}}(y_{\tilde{b}})} |h_{\tilde{b}}''(y_{\tilde{b}})|^{-1/2} N^{-1/2} \left[\int_0^{\delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2} \sqrt{N}} e^{-\frac{1}{2} s^2} \left| e^{-\mathbf{i} \frac{1}{6} h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} s/\sqrt{N}))} [-h_{\tilde{b}}''(y_{\tilde{b}})]^{-3/2} s^3/\sqrt{N} - 1 \right| \frac{ds}{2\pi} \right. \\
&\quad \left. + \int_{-\delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2} \sqrt{N}}^0 e^{-\frac{1}{2} s^2} \left| e^{-\mathbf{i} \frac{1}{6} h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + 2\pi\mathbf{i} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} s/\sqrt{N}))} [-h_{\tilde{b}}''(y_{\tilde{b}})]^{-3/2} s^3/\sqrt{N} - 1 \right| \frac{ds}{2\pi} \right].
\end{aligned}$$

If $\tilde{b} \in [\tau, 1 - \tau]$, then uniformly in $s \in [0, \delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2} \sqrt{N}]$ we have

$$\delta_{\tilde{b}}^3 h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} s/\sqrt{N})) N = o(1).$$

Otherwise, by (3.19), we have uniformly in complex $t \in B(0, \delta_{\tilde{b}})$ that

$$|h_{\tilde{b}}'''(y_{\tilde{b}} + t)| \leq C_1' \tilde{b} \text{ if } \tilde{b} \in [0, \tau] \quad |h_{\tilde{b}}'''(y_{\tilde{b}} + t)| \leq C_1'(1 - \tilde{b}) \text{ if } \tilde{b} \in [1 - \tau, 1],$$

meaning that uniformly in $s \in [0, \delta_{\tilde{b}} |h_{\tilde{b}}''(y_{\tilde{b}})|^{1/2} \sqrt{N}]$ we have for $\tilde{b} \in [0, \tau]$ that

$$|\delta_{\tilde{b}}^3 h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} s/\sqrt{N})) N| \leq C_1' N^{-6/5} |h_{\tilde{b}}''(y_{\tilde{b}})|^{-6/5} \tilde{b} N = O((\tilde{b} N)^{-1/5}).$$

A similar argument shows this bound for $\tilde{b} \in [1 - \tau, 1]$ and for $|\delta_{\tilde{b}}^3 h_{\tilde{b}}'''(\xi_{\tilde{b}}(y_{\tilde{b}} + 2\pi\mathbf{i} + \mathbf{i}[-h_{\tilde{b}}''(y_{\tilde{b}})]^{-1/2} s/\sqrt{N})) N|$. We conclude that

$$|K_3(\tilde{b})| = O((\tilde{b} N)^{-1/5}) |K_1(\tilde{b})| = O(k_N^{-1/5}) |K_1(\tilde{b})| = |K_1(\tilde{b})| \cdot o(1).$$

We conclude that

$$K(\tilde{b}) = \frac{1}{\sqrt{2\pi N h_{\tilde{b}}''(y_{\tilde{b}})}} e^{-N h_{\tilde{b}}(y_{\tilde{b}})} (1 + o(1)).$$

Putting these computations together, we conclude that

$$(3.27) \quad I_1(\tilde{a}, \tilde{b}) = \frac{-\mathbf{i}}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \frac{1}{2\pi N \sqrt{h_{\tilde{a}}''(y_{\tilde{a}})} \sqrt{h_{\tilde{b}}''(y_{\tilde{b}})}} e^{N[h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{b}}(y_{\tilde{b}})]} (1 + o(1)).$$

We now bound $I_2(\tilde{a}, \tilde{b})$, $I_3(\tilde{a}, \tilde{b})$, and $I_4(\tilde{a}, \tilde{b})$. For $I_2(\tilde{a}, \tilde{b})$, by (3.22) and (3.14) and the construction of $\Gamma_{\tilde{a}}$, on $\Gamma_{\tilde{a}}^2$ we have

$$\left| \exp \left(N(h_{\tilde{a}}(\tilde{z}) - h_{\tilde{a}}(y_{\tilde{a}})) \right) \right| < \exp \left(-\frac{1}{4} |h_{\tilde{a}}''(y_{\tilde{a}})| \delta_{\tilde{a}}^2 N \right) \leq e^{-CN^{1/5}}$$

for some $C > 0$, which implies by the fact that $\Gamma_{\tilde{a}}$ and $\Sigma_{\tilde{b}}$ have bounded length that

$$|I_2(\tilde{a}, \tilde{b})| \leq e^{-CN^{1/5}} \exp \left(N(h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{b}}(y_{\tilde{b}})) \right) = |I_1(\tilde{a}, \tilde{b})| \cdot o(1)$$

uniformly in $\tilde{a}, \tilde{b}$. Similarly, for $I_3(\tilde{a}, \tilde{b})$, by (3.23), (3.16), (3.17), and the construction of $\Gamma_{\tilde{b}}$, on $\Gamma_{\tilde{b}}^2$ we have for large enough N that

$$\left| \exp \left(N(h_{\tilde{b}}(y_{\tilde{b}}) - h_{\tilde{b}}(\tilde{w})) \right) \right| < \exp \left(-\frac{1}{4} |h_{\tilde{b}}''(y_{\tilde{b}})| \delta_{\tilde{b}}^2 N \right) = e^{-[N|h_{\tilde{b}}''(y_{\tilde{b}})|]^{1/5}/4}.$$

This implies by the fact that $\Gamma_{\tilde{a}}$ and $\Sigma_{\tilde{b}}$ have bounded length that

$$|I_3(\tilde{a}, \tilde{b})| \leq e^{-[N|h_{\tilde{b}}''(y_{\tilde{b}})|]^{1/5}/4} e^{Nh_{\tilde{a}}(y_{\tilde{a}})} \oint_{\Gamma_{\tilde{a}}} \frac{d\tilde{z}}{2\pi i} \oint_{\Sigma_{\tilde{b}}} \frac{d\tilde{w}}{2\pi i} \left| \frac{1}{1 - e^{\tilde{w} - \tilde{z}}} \exp \left(Nh_{\tilde{a}}(\tilde{z}) \right) \right| = o(1) \cdot |I_1(\tilde{a}, \tilde{b})|,$$

where in the second step we apply a standard steepest descent analysis in $\tilde{z}$.

For $I_4(\tilde{a}, \tilde{b})$, for $\tilde{z} \in \Gamma_{\tilde{a}}^1$ and $\tilde{w} \in \Gamma_{\tilde{b}}^1$, since $\delta_{\tilde{a}} = o(1)$ and $\delta_{\tilde{b}} = o(1)$, we have for large enough N that

$$\left| \frac{1}{1 - e^{\tilde{w} - \tilde{z}}} - \frac{1}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \right| \leq 3 \left| \frac{e^{y_{\tilde{b}} - y_{\tilde{a}}}}{(1 - e^{y_{\tilde{b}} - y_{\tilde{a}}})^2} \right| (\delta_{\tilde{a}} + \delta_{\tilde{b}}) = o(1)$$

uniformly in $\tilde{b}$. We conclude that

$$(3.28) \quad |I_4(\tilde{a}, \tilde{b})| = o(1) \left[\oint_{\Gamma_{\tilde{a}}^1} e^{Nh_{\tilde{a}}(\tilde{z})} \frac{d\tilde{z}}{2\pi i} \right] \left[\oint_{\Gamma_{\tilde{b}}^1} e^{-Nh_{\tilde{b}}(\tilde{w})} \frac{d\tilde{w}}{2\pi i} \right] = o(1) \cdot |I_1(\tilde{a}, \tilde{b})|.$$

Combining our expressions for $I_1(\tilde{a}, \tilde{b})$, $I_2(\tilde{a}, \tilde{b})$, $I_3(\tilde{a}, \tilde{b})$, and $I_4(\tilde{a}, \tilde{b})$ yields (3.24).

Case 1, Step 3. We now account for $R(\tilde{a}, \tilde{b})$, which is given by

$$R(\tilde{a}, \tilde{b}) = \frac{1}{(\tilde{a} - \tilde{b})N} \left[e^{N(\tilde{a} - \tilde{b})c_1} - e^{N(\tilde{a} - \tilde{b})c_2} \right],$$

where c_1, c_2 are the endpoints of Ξ which lie on $\Gamma_{\tilde{a}} \cap \Sigma_{\tilde{b}}$. For large enough N , these endpoints must lie on $\Gamma_{\tilde{a}}^2 \cap \Sigma_{\tilde{b}}^2$, so we see by (3.22) that

$$\Re[(\tilde{a} - \tilde{b})c_i] = \Re[h_{\tilde{a}}(c_i) - h_{\tilde{b}}(c_i)] < \Re[h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{b}}(y_{\tilde{b}})] - \frac{1}{4} |h_{\tilde{a}}''(y_{\tilde{a}})| \delta_{\tilde{a}}^2,$$

which implies that

$$|R(\tilde{a}, \tilde{b})| < e^{-\frac{1}{4} |h_{\tilde{a}}''(y_{\tilde{a}})| \delta_{\tilde{a}}^2 N} \left| \exp \left(N(h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{b}}(y_{\tilde{b}})) \right) \right| = o(1) \cdot |I'(\tilde{a}, \tilde{b})|.$$

We conclude finally that

$$I(\tilde{a}, \tilde{b}) = \frac{-i}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \frac{1}{2\pi N \sqrt{h_{\tilde{a}}''(y_{\tilde{a}})} \sqrt{h_{\tilde{b}}''(y_{\tilde{b}})}} e^{N[h_{\tilde{a}}(y_{\tilde{a}}) - h_{\tilde{b}}(y_{\tilde{b}})]} (1 + o(1)).$$

Case 1, Step 4. It remains to analyze the prefactor $\frac{J(\tilde{a}, \tilde{b})}{I(\tilde{a}, \tilde{b})}$. Applying Stirling's approximation, we see that

$$\Gamma(N - \tilde{b}N) \Gamma(\tilde{b}N + 1) = 2\pi \sqrt{\frac{\tilde{b}}{1 - \tilde{b}}} N^N e^{-N} (1 - \tilde{b})^{(1 - \tilde{b})N} \tilde{b}^{\tilde{b}N} [1 + o(1)].$$

For each $\tilde{a} \in \mathbb{C}$, choose one of two branches of the logarithm such that the branch line avoids the set $[\tilde{a} - 1, \tilde{a}]$ in the complex plane. For this branch, we see that

$$\begin{aligned} \log \frac{\Gamma(\tilde{a}N - N + 1)}{\Gamma(\tilde{a}N + 1)} &= - \sum_{i=0}^{N-1} \log(\tilde{a}N - i) \\ &= -N \log N - \sum_{i=0}^{N-1} \log(\tilde{a} - i/N) \\ &= -N \log N - \frac{1}{2} [\log(\tilde{a}) - \log(\tilde{a} - 1)] - N \int_0^1 \log(\tilde{a} - x) dx + O(N^{-1}) \\ &= -N \log N + N - \frac{1}{2} [\log(\tilde{a}) - \log(\tilde{a} - 1)] - (1 - \tilde{a}) \log(\tilde{a} - 1)N - \tilde{a} \log(\tilde{a})N + O(N^{-1}) \end{aligned}$$

uniformly in $\tilde{a}$, where we apply the trapezoid rule. Exponentiating, we find that

$$\frac{\Gamma(\tilde{a}N - N + 1)}{\Gamma(\tilde{a}N + 1)} = N^{-N} e^N \sqrt{\frac{\tilde{a}-1}{\tilde{a}}} \left(\frac{\tilde{a}-1}{\tilde{a}}\right)^{\tilde{a}N} \frac{1}{(\tilde{a}-1)^N} [1 + o(1)],$$

where we take the standard branch of the square root in $\sqrt{\frac{\tilde{a}-1}{\tilde{a}}}$ and note that $\frac{\tilde{a}-1}{\tilde{a}}$ does not enclose the origin for $\tilde{a} \in \Upsilon_\varepsilon$ and therefore has holomorphic square root. We thus conclude that

$$(-1)^{(1-\tilde{b})N+1} \frac{\Gamma(N - \tilde{b}N) \Gamma(\tilde{b}N + 1) \Gamma(\tilde{a}N - N + 1) (\tilde{a} - \tilde{b})^N}{\Gamma(\tilde{a}N + 1)} = -2\pi(\tilde{a} - \tilde{b})^N \frac{\sqrt{\frac{\tilde{b}}{1-\tilde{b}}} (\tilde{b} - 1)^{(1-\tilde{b})N} \tilde{b}^{\tilde{b}N}}{\sqrt{\frac{\tilde{a}}{\tilde{a}-1}} (1 - \tilde{a}^{-1})^{-\tilde{a}N} (\tilde{a} - 1)^N} [1 + o(1)]$$

and hence

$$\tilde{B}_N(\tilde{a}, \tilde{b}) = \frac{1}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \frac{\tilde{a} - \tilde{b}}{\sqrt{h''_{\tilde{a}}(y_{\tilde{a}})} \sqrt{h''_{\tilde{b}}(y_{\tilde{b}})}} \frac{\sqrt{\frac{\tilde{b}}{1-\tilde{b}}} (\tilde{b} - 1)^{(1-\tilde{b})N} \tilde{b}^{\tilde{b}N} e^{-N h_{\tilde{b}}(y_{\tilde{b}})}}{\sqrt{\frac{\tilde{a}}{\tilde{a}-1}} (1 - \tilde{a}^{-1})^{-\tilde{a}N} (\tilde{a} - 1)^N e^{-N h_{\tilde{a}}(y_{\tilde{a}})}} [1 + o(1)],$$

where the $o(1)$ is uniform in $\tilde{a}$ and $\tilde{b}$. Substituting $y_{\tilde{a}} = -\log M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1)$ and $y_{\tilde{b}} = -\log M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)$ and noting that

$$(1 - c) \log(c - 1) + c \log(c) - h_c(y_c) = \tilde{\Psi}_{\rho_N}(c),$$

we find that

$$\begin{aligned} \tilde{B}_N(\tilde{a}, \tilde{b}) &= \frac{M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)}{M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1)} \\ &\quad \frac{\tilde{a} - \tilde{b}}{\sqrt{M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1) M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1))} \sqrt{M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1) M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1))}} \frac{\sqrt{\frac{\tilde{b}}{1-\tilde{b}}} e^{N \tilde{\Psi}_{\tilde{\rho}_N}(\tilde{b})}}{\sqrt{\frac{\tilde{a}}{\tilde{a}-1}} e^{N \tilde{\Psi}_{\tilde{\rho}_N}(\tilde{a})}} [1 + o(1)]. \end{aligned}$$

Finally, since $\tilde{b} \in (0, 1)$, for the standard branch of the square root we have

$$\frac{M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)}{\sqrt{M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1) M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1))}} \cdot \sqrt{\frac{\tilde{b}}{1-\tilde{b}}} = -\frac{\sqrt{\frac{\tilde{b}}{1-\tilde{b}}} M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1))}} = -\frac{\sqrt{S_{\tilde{\rho}_N}(\tilde{b} - 1)}}{\sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1))}}.$$

Similarly, for $\tilde{a} \in \Upsilon_\varepsilon$ we see that

$$\sqrt{M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1) M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1))} \cdot \sqrt{\frac{\tilde{a}}{\tilde{a}-1}} = \sqrt{S_{\tilde{\rho}_N}(\tilde{a} - 1)} \cdot \sqrt{M'_{\tilde{\rho}_N}(M_{\tilde{\rho}_N}^{-1}(\tilde{a} - 1))},$$

where each of the arguments for the square roots on the left lie in the right half plane, meaning that this equality holds for the standard square root. Substituting these equalities yields the desired result.

Case 2: $0 \leq \tilde{b} \leq k_N/N$: Contracting the w -contour through 0, we find that

$$I(\tilde{a}, \tilde{b}) = e^{\sum_i \lambda_i^N} \sum_{k=0}^{\tilde{b}N} \oint_{\{e^{\lambda_i^N}\}} \frac{dz}{2\pi i} (-1)^{N-k} e_k(e^{-\lambda_i^N}) z^{(\tilde{a}-\tilde{b})N+k} \prod_{i=1}^N \frac{1}{z - e^{\lambda_i^N}},$$

where e_k is the k^{th} elementary symmetric polynomial. Because λ_i^N are all contained in the fixed finite interval I of Assumption 3.3, we have a uniform upper bound $e^{\lambda_i^N} < M$. This implies that

$$\left| \frac{p_1(e^{-\lambda_i^N})^k}{k!} - e_k(e^{-\lambda_i^N}) \right| \leq \frac{M^k}{k!} (N^k - N(N-1) \cdots (N-k+1)) = \frac{N^k M^k}{k!} \left(1 - \prod_{i=0}^{k-1} (1 - i/N)\right).$$

Notice now that

$$\log \prod_{i=0}^{k-1} (1 - i/N) \geq N \int_0^{\frac{k-1}{N}} \log(1-x) dx = -(k-1) + (k-N-1) \log \frac{N-k+1}{N},$$

which implies using the fact that $1 - \varepsilon \leq e^{-\varepsilon}$ that

$$\prod_{i=0}^{k-1} (1 - i/N) \geq e^{1-k} \left(1 - \frac{k-1}{N}\right)^{k-N-1} \geq \exp\left(1 - k - \frac{(k-1)(k-N-1)}{N}\right) = e^{-\frac{(k-1)^2}{N}}.$$

These imply that if $M^k \frac{k^2}{N} = o(1)$, then uniformly in k we have

$$e_k(e^{-\lambda_i^N}) = \frac{p_1(e^{-\lambda_i^N})^k}{k!} [1 + o(1)].$$

By the bound on $0 \leq \tilde{b} \leq k_N/N$ and (3.20), we find that

$$\begin{aligned} I(\tilde{a}, \tilde{b}) &= e^{\sum_i \lambda_i^N} \sum_{k=0}^{\tilde{b}N} \oint_{\{e^{\lambda_i^N}\}} \frac{dz}{2\pi i} (-1)^{N-k} \frac{N^k}{k!} \left(\frac{p_1(e^{-\lambda_i^N})}{N} \right)^k z^{(\tilde{a}-\tilde{b})N+k} \prod_{i=1}^N \frac{1}{z - e^{\lambda_i^N}} [1 + o(1)] \\ &= e^{\sum_i \lambda_i^N} \oint_{\Gamma_{\tilde{a}}} \frac{d\tilde{z}}{2\pi i} (-1)^{N-\tilde{b}N} \frac{N^{\tilde{b}N}}{(\tilde{b}N)!} \left(\frac{p_1(e^{-\lambda_i^N})}{N} \right)^{\tilde{b}N} e^{Nh_{\tilde{a}}(\tilde{z})} [1 + o(1)] \\ &= -e^{\sum_i \lambda_i^N} (-1)^{N-\tilde{b}N} \frac{N^{\tilde{b}N}}{(\tilde{b}N)!} \left(\frac{p_1(e^{-\lambda_i^N})}{N} \right)^{\tilde{b}N} \frac{1}{\sqrt{2\pi N h''_{\tilde{a}}(y_{\tilde{a}})}} e^{Nh_{\tilde{a}}(y_{\tilde{a}})} [1 + o(1)], \end{aligned}$$

where the second step holds uniformly in $\tilde{b}$ because we replace the sum by the asymptotically largest term and the penultimate step is by a steepest descent analysis in $\tilde{z}$ similar to Case 1. In this setting, we conclude as before that

$$\frac{\Gamma(\tilde{a}N - N + 1)}{\Gamma(\tilde{a}N + 1)} = N^{-N} e^N \sqrt{\frac{\tilde{a} - 1}{\tilde{a}}} \frac{1}{(\tilde{a} - 1)^{(1-\tilde{a})N} \tilde{a}^{\tilde{a}N}} [1 + o(1)]$$

and that

$$\Gamma(N - \tilde{b}N) = \sqrt{\frac{2\pi}{(1-\tilde{b})N}} N^{N-\tilde{b}N} (1-\tilde{b})^{(1-\tilde{b})N} e^{-N+\tilde{b}N} [1 + o(1)].$$

We find that

$$\tilde{B}_N(\tilde{a}, \tilde{b}) = -\frac{\tilde{a} - \tilde{b}}{\sqrt{h''_{\tilde{a}}(y_{\tilde{a}})}} \left(\frac{p_1(e^{-\lambda_i^N})}{N} \right)^{\tilde{b}N} \frac{e^{\sum_i \lambda_i^N} e^{\tilde{b}N} \frac{1}{\sqrt{1-\tilde{b}}} (1-\tilde{b})^{(1-\tilde{b})N}}{\sqrt{\frac{\tilde{a}}{\tilde{a}-1}} e^{N\tilde{\Psi}_{\rho_N}(\tilde{a})}} [1 + o(1)]$$

uniformly in $\tilde{a}$ and $\tilde{b}$. Notice now that

$$M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)^{-1} = -\frac{\tilde{b}}{p_1(e^{-\lambda_i^N})/N} [1 + O(k_N/N)] \quad \text{and} \quad \frac{1}{N} p_1(e^{-\lambda_i^N}) = \int e^{-s} d\rho_N(s).$$

Noting that $k_N^2/N = o(1)$, we find that

$$\begin{aligned} \exp\left(Nh_{\tilde{b}}(y_{\tilde{b}})\right) &= (-1)^N M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)^{-\tilde{b}N} e^{-\sum_i \lambda_i^N} \exp\left(-N \int \log(1 - M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)^{-1} e^{-s}) d\rho_N(s)\right) \\ &= (-1)^{(1-\tilde{b})N} \tilde{b}^{\tilde{b}N} (p_1(e^{-\lambda_i^N})/N)^{-\tilde{b}N} e^{-\sum_i \lambda_i^N} \exp\left(N M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)^{-1} \int e^{-s} d\rho_N(s) + O(N\tilde{b}^2)\right) [1 + o(1)] \\ &= (-1)^{(1-\tilde{b})N} \tilde{b}^{\tilde{b}N} (p_1(e^{-\lambda_i^N})/N)^{-\tilde{b}N} e^{-\sum_i \lambda_i^N} e^{-\tilde{b}N} [1 + o(1)] \end{aligned}$$

and that

$$\begin{aligned} h''_{\tilde{b}}(y_{\tilde{b}}) &= \int \frac{e^s M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)}{(1 - e^s M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1))^2} d\rho_N(s) = -\tilde{b} + O(\tilde{b}^2) = -\tilde{b} [1 + o(1)] \\ e^{y_{\tilde{b}} - y_{\tilde{a}}} &= M_{\tilde{\rho}_N}^{-1}(\tilde{b} - 1)^{-1} e^{-y_{\tilde{a}}} = o(1). \end{aligned}$$

uniformly in $\tilde{b}$. Together, these imply that

$$e^{N\tilde{\Psi}_{\rho_N}(\tilde{b})} \frac{\tilde{b}^{1/2}}{\sqrt{h''_{\tilde{b}}(y_{\tilde{b}})}} \frac{1}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} = -\mathbf{i}(1 - \tilde{b})^{(1-\tilde{b})N} (p_1(e^{-\lambda_i^N})/N)^{\tilde{b}N} e^{\sum_i \lambda_i^N} e^{\tilde{b}N} [1 + o(1)]$$

Substituting in, we find that

$$\tilde{B}_N(\tilde{a}, \tilde{b}) = \frac{\mathbf{i}}{1 - e^{y_{\tilde{b}} - y_{\tilde{a}}}} \frac{\tilde{a} - \tilde{b}}{\sqrt{h''_{\tilde{a}}(y_{\tilde{a}})} \sqrt{h''_{\tilde{b}}(y_{\tilde{b}})}} \frac{\sqrt{\frac{\tilde{b}}{1-\tilde{b}}} e^{N\tilde{\Psi}_{\rho_N}(\tilde{b})}}{\sqrt{\frac{\tilde{a}}{\tilde{a}-1}} e^{N\tilde{\Psi}_{\rho_N}(\tilde{a})}} [1 + o(1)],$$

which simplifies to the desired expression.

Case 3: $1 - k_N/N \leq \tilde{b} < 1$: After deforming the w -contour to ∞ , we obtain

$$I(\tilde{a}, \tilde{b}) = \sum_{k=0}^{(1-\tilde{b})N} \oint_{e^{\lambda_i^N}} \frac{dz}{2\pi\mathbf{i}} (-1)^{k+1} e_k(e^{\lambda_i^N}) z^{(\tilde{a}+1-\tilde{b})N-k-1} \prod_{i=1}^N \frac{1}{z - e^{\lambda_i^N}}.$$

The remainder of the analysis then proceeds similarly to Case 2. □

3.5. Proof of Theorem 3.4. By Proposition 3.1, we have that

$$\tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) = J_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) I_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k)$$

for

$$J_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) := \frac{\prod_{m \neq l} (\tilde{a}_m N - \tilde{b}_l N)}{\Delta(\tilde{a}_m N) \Delta(\tilde{b}_m N)} = \frac{\prod_{m \neq l} (\tilde{a}_m - \tilde{b}_l)}{\Delta(\tilde{a}) \Delta(\tilde{b})}$$

$$I_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) := (-1)^{\frac{k(k-1)}{2}} \det \left(\frac{(\tilde{a}_i - \tilde{b}_i)N}{(\tilde{a}_i N - \tilde{b}_j N)} \frac{\mathcal{B}(\mu_{\tilde{a}_i N, \tilde{b}_j N}, \lambda^N)}{\mathcal{B}(\rho, \lambda^N)} \right)_{i,j=1}^k.$$

By Theorem 3.9, we have

$$\begin{aligned} I_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k) &= (-1)^{\frac{k(k-1)}{2}} \det \left(\frac{\tilde{a}_i - \tilde{b}_i}{\tilde{a}_i - \tilde{b}_j} \tilde{B}_N(\tilde{a}_i, \tilde{b}_j) \right)_{i,j=1}^k \\ &= (-1)^{\frac{k(k-1)}{2}} \det \left(\frac{1}{M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1)} \right. \\ &\quad \left. \frac{\tilde{a}_i - \tilde{b}_i}{\sqrt{M_{\tilde{\rho}_N}'(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1)) M_{\tilde{\rho}_N}'(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1))}} \frac{\sqrt{S_{\tilde{\rho}_N}(\tilde{b}_j - 1)} e^{N\tilde{\Psi}_{\rho_N}(\tilde{b}_j)}}{\sqrt{S_{\tilde{\rho}_N}(\tilde{a}_i - 1)} e^{N\tilde{\Psi}_{\rho_N}(\tilde{a}_i)}} [1 + o(1)] \right)_{i,j=1}^k \\ &= (-1)^{\frac{k(k-1)}{2}} \det \left(\frac{1}{M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1)} [1 + o(1)] \right)_{i,j=1}^k \\ &\quad \prod_{i=1}^k \left[\frac{(\tilde{a}_i - \tilde{b}_i) \sqrt{S_{\tilde{\rho}_N}(\tilde{b}_i - 1)} e^{N\tilde{\Psi}_{\rho_N}(\tilde{b}_i)}}{\sqrt{M_{\tilde{\rho}_N}'(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1)) M_{\tilde{\rho}_N}'(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))} \sqrt{S_{\tilde{\rho}_N}(\tilde{a}_i - 1)} e^{N\tilde{\Psi}_{\rho_N}(\tilde{a}_i)}} \right] \\ &= \frac{\Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1)) \Delta(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))}{\prod_{i,j} (M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1) - M_{\tilde{\rho}_N}^{-1}(\tilde{b}_j - 1))} \\ &\quad \prod_{i=1}^k \left[\frac{(\tilde{a}_i - \tilde{b}_i) \sqrt{S_{\tilde{\rho}_N}(\tilde{b}_i - 1)} e^{N\tilde{\Psi}_{\rho_N}(\tilde{b}_i)}}{\sqrt{M_{\tilde{\rho}_N}'(M_{\tilde{\rho}_N}^{-1}(\tilde{a}_i - 1)) M_{\tilde{\rho}_N}'(M_{\tilde{\rho}_N}^{-1}(\tilde{b}_i - 1))} \sqrt{S_{\tilde{\rho}_N}(\tilde{a}_i - 1)} e^{N\tilde{\Psi}_{\rho_N}(\tilde{a}_i)}} \right] [1 + o(1)], \end{aligned}$$

where uniformity in $\tilde{b}_1, \dots, \tilde{b}_k$ follows from the corresponding uniformity in Theorem 3.9 and where in the last step we use the Cauchy determinant formula. Combining these asymptotics yields the desired.

4. PRODUCTS OF RANDOM MATRICES

4.1. Multivariate Bessel generating functions for matrix products. For a $N \times N$ Hermitian positive-definite random matrix X with measure $d\sigma(X)$, let $\mu = (\mu_1 \geq \dots \geq \mu_N > 0)$ be its spectrum, which inherits a measure $d\sigma(\mu)$. If the log-spectral measure of X is ρ -smooth for $\rho = (N-1, \dots, 1, 0)$, then its multivariate Bessel generating function with respect to ρ is given by

$$(4.1) \quad \phi_X(s) := \int \frac{\mathcal{B}(s, \log \mu)}{\mathcal{B}(\rho, \log \mu)} d\sigma(\mu).$$

In this section, we will only consider multivariate Bessel generating functions with respect to ρ , so we will omit the N -tuple χ of (2.2) from the notations. In the following two lemmas, we prove the key property that the multivariate Bessel generating function is multiplicative over products of matrices.

Lemma 4.1. When s is an integral signature such that $s - \rho$ is a partition, the multivariate Bessel generating function of a random matrix X with smooth log-spectral measure takes the form

$$\phi_X(s) = \frac{1}{\dim L_{s-\rho}} \int \chi_{s-\rho}(X) d\sigma(X),$$

where $L_{s-\rho}$ is the highest weight representation of $\mathfrak{gl}_N$ corresponding to $s - \rho$ and $\chi_{s-\rho}$ is its character.

Proof. By the Weyl character formula, for a matrix X with spectrum μ , we have

$$\chi_{s-\rho}(X) = \frac{\det(\mu_i^{s_j})_{i,j=1}^N}{\Delta(\mu)}.$$

Combining this with (2.1) and (4.1), we find that

$$\phi_X(s) = \int \frac{\det(\mu_i^{s_j}) \Delta(\rho)}{\Delta(\mu) \Delta(s)} d\sigma(\mu) = \frac{1}{\dim L_{s-\rho}} \int \chi_{s-\rho}(X) d\sigma(X). \quad \square$$

Lemma 4.2. Let $X_1 = Y_1^* Y_1$ and $X_2 = Y_2^* Y_2$ be Hermitian random matrices with spectral measures $d\sigma_1(X_1)$ and $d\sigma_2(X_2)$, and let $X_3 = (Y_1 Y_2)^* (Y_1 Y_2)$. If the distribution of X_2 is unitarily invariant and the log-spectral measures of X_1 and X_2 are smooth, then we have the identity

$$\phi_{X_3}(s) = \phi_{X_1}(s) \phi_{X_2}(s)$$

of multivariate Bessel generating functions.

Proof. Both sides satisfy the exponential growth conditions of Carlson's theorem (Theorem 3.7), so it suffices to check this for integral signatures s such that $s - \rho$ is a partition. Noting that X_3 has the same spectrum as $X_1 X_2$ and the functional relation

$$\int_U \chi_{s-\rho}(X_1 U X_2 U^*) d\text{Haar}_U = \frac{1}{\dim L_{s-\rho}} \chi_{s-\rho}(X_1) \chi_{s-\rho}(X_2)$$

from [Mac95, Lemma VII.4.2] and [For10, Proposition 13.4.2], we may compute by Lemma 4.1 that

$$\begin{aligned} \phi_{X_3}(s) &= \frac{1}{\dim L_{s-\rho}} \iint \chi_{s-\rho}(X_1 X_2) d\sigma_1(X_1) d\sigma_2(X_2) \\ &= \frac{1}{\dim L_{s-\rho}} \int_U \iint \chi_{s-\rho}(X_1 U X_2 U^*) d\sigma_1(X_1) d\sigma_2(X_2) d\text{Haar}_U \\ &= \frac{1}{(\dim L_{s-\rho})^2} \int \chi_{s-\rho}(X_1) d\sigma_1(X_1) \cdot \int \chi_{s-\rho}(X_2) d\sigma_2(X_2) \\ &= \phi_{X_1}(s) \phi_{X_2}(s), \end{aligned}$$

where we apply unitary invariance in the second equality and note that the log-spectral measure of X_3 is smooth by this expression for the defining integral of its multivariate Bessel generating function. $\square$

4.2. LLN-appropriate and CLT-appropriate measures. For each N , let $Y_N^1, \dots, Y_N^k$ be $N \times N$ random matrices which are right unitarily invariant, and let $X_N^i := (Y_N^i)^* Y_N^i$. We will study the log-spectral measure of

$$X_N := (Y_N^1 \cdots Y_N^k)^* (Y_N^1 \cdots Y_N^k),$$

which is a $N \times N$ positive definite Hermitian random matrix. For this, notice that the multivariate Bessel generating function of such measures takes a simple form.

Corollary 4.3. The multivariate Bessel generating function of the log-spectral measure of X_N is given by

$$\phi_{X_N}(s) = \prod_{i=1}^k \phi_{X_N^i}(s).$$

Proof. This follows by iteratively applying Lemma 4.2. $\square$

Lemma 4.4. If the log-spectral measures of $X_N^1, \dots, X_N^k$ are LLN-appropriate (CLT-appropriate) for some $\Psi_i(u)$ ($\Lambda_i(u, w)$), then so is the log-spectral measure of X_N for

$$\Psi(u) := \sum_{i=1}^k \Psi_i(u) \quad \text{and} \quad \Lambda(u, w) := \sum_{i=1}^k \Lambda_i(u, w).$$

Proof. This follows directly from Corollary 4.3. $\square$

Remark. Corollary 4.3 reflects the fact that the log-spectral measure of X_N depends only on the distributions of $X_N^i = (Y_N^i)^* Y_N^i$. In what follows, we will state our results in terms of the distribution of X_N^i instead of Y_N^i , meaning that they will hold for any distribution on Y_N^i for which $X_N^i \stackrel{d}{=} (Y_N^i)^* Y_N^i$.

We now give several situations in which the spectral measure of a single random matrix is LLN-appropriate and CLT-appropriate.

Theorem 4.5. Let X_N be a sequence of unitarily invariant $N \times N$ positive-definite Hermitian random matrices with spectrum λ^N . If the empirical measure of λ^N has support contained within a fixed finite interval and converges weakly to a compactly supported measure $d\rho$, then the log-spectral measure of X_N is CLT-appropriate with

$$\begin{aligned} \Psi(u) &= -\log S_{\tilde{\rho}}(u-1) \\ \Lambda(u, w) &= \frac{1}{M_{\tilde{\rho}}'(M_{\tilde{\rho}}^{-1}(u-1))M_{\tilde{\rho}}'(M_{\tilde{\rho}}^{-1}(w-1))(M_{\tilde{\rho}}^{-1}(u-1) - M_{\tilde{\rho}}^{-1}(w-1))^2} - \frac{1}{(u-w)^2}. \end{aligned}$$

Proof. By definition, we have that

$$\phi_{X_N}(s) = \frac{\mathcal{B}(s, \log \lambda^N)}{\mathcal{B}(\rho, \log \lambda^N)},$$

which means that for $s = \mu_{\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k}$ from (3.3), we have that

$$\phi_{X_N}(\mu_{\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k}) = \tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k).$$

For $I = \{i_1, \dots, i_k\}$, notice by Theorem 3.4 and the fact that uniformly convergent sequences of holomorphic functions can be differentiated that

$$\frac{1}{N} \partial_{r_{i_1}} \log \phi_{X_N}^I(rN) = \frac{1}{N} \partial_{\tilde{a}_1} [\log \tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k)] \Big|_{\tilde{a}_m=r_{i_m}, \tilde{b}_m=\frac{N-i_m}{N}} = -\tilde{\Psi}'_{\rho_N}(r_{i_1}) + o(1)$$

uniformly in i , I , and r^I in a neighborhood of $[0, 1]^{|I|}$, where we recall from (3.4) that

$$\tilde{\Psi}_{\rho_N}(r) = r \log S_{\tilde{\rho}_N}(r-1) + \log(r-1) + \int \log(M_{\tilde{\rho}_N}^{-1}(r-1)^{-1} - e^s) d\rho(s).$$

Taking the limit $N \rightarrow \infty$ and applying Lemma 3.10, we find that

$$\begin{aligned} \lim_{N \rightarrow \infty} -\tilde{\Psi}'_{\rho_N}(u) &= -\log S_{\tilde{\rho}}(u-1) - u \frac{S'_{\tilde{\rho}}(u-1)}{S_{\tilde{\rho}}(u-1)} - \frac{1}{u-1} - \int \frac{\frac{\partial}{\partial u} M_{\tilde{\rho}}^{-1}(u-1)^{-1}}{M_{\tilde{\rho}}^{-1}(u-1)^{-1} - e^s} d\rho(s) \\ &= -\log S_{\tilde{\rho}}(u-1) - u \frac{S'_{\tilde{\rho}}(u-1)}{S_{\tilde{\rho}}(u-1)} - \frac{1}{u-1} - u M_{\tilde{\rho}}^{-1}(u-1) \frac{\partial}{\partial u} M_{\tilde{\rho}}^{-1}(u-1)^{-1} \\ &= -\log S_{\tilde{\rho}}(u-1). \end{aligned}$$

Again by Theorem 3.4, we have that

$$\begin{aligned} \partial_{r_{i_1}} \partial_{r_{i_2}} \log \phi_{X_N}^I(rN) &= \partial_{\tilde{a}_1} \partial_{\tilde{a}_2} [\log \tilde{B}_N(\tilde{a}_1, \dots, \tilde{a}_k; \tilde{b}_1, \dots, \tilde{b}_k)] \Big|_{\tilde{a}_m=r_{i_m}, \tilde{b}_m=\frac{N-i_m}{N}} \\ &= F_N^{(1,1)}(r_{i_1}, r_{i_2}) + o(1) \end{aligned}$$

uniformly in i_1, i_2, I , and r^I in a neighborhood of $[0, 1]^{|I|}$ for

$$F_N(u, w) = \log(M_{\rho_N}^{-1}(u-1) - M_{\rho_N}^{-1}(w-1)) - \log(u-w) = F(u, w) + o(1).$$

Noting that $\Lambda(u, w) = F^{(1,1)}(u, w)$, we conclude that X_N is CLT-appropriate for $\Psi(u)$ and $\Lambda(u, w)$. $\square$

We now consider the Jacobi and Wishart ensembles. At $\beta = 2$, the Jacobi ensemble with parameters α and $R \geq N$ is the process on N points in $[0, 1]$ with density proportional to

$$\prod_{i < j} (\lambda_i - \lambda_j)^2 \prod_{i=1}^N \lambda_i^{\alpha-1} (1 - \lambda_i)^{R-N}.$$

Let Y be a $N \times (\alpha + N - 1)$ submatrix of a Haar distributed matrix from the unitary group $U_{\alpha+R+N-1}$; then by [For10, Proposition 3.8.2], the Jacobi ensemble is the distribution of the eigenvalues of Y^*Y . Alternatively, as discussed in [For10, Proposition 3.6.1], it corresponds to the spectral distribution of the matrix

$$(X^{AR})^* X^{AR} \left((X^{AR})^* X^{AR} + (Y^{NR})^* Y^{NR} \right)^{-1},$$

where X^{AR} and Y^{NR} are $A \times R$ and $N \times R$ matrices of i.i.d. complex Gaussians and $A = \alpha + R - 1$. We now show that this ensemble fits into our framework, for which we must introduce the Heckman-Opdam hypergeometric function at $\beta = 2$.² It and its normalized version are given by

$$\mathcal{F}_r(s) := \Delta(r) \mathcal{B}(r, s) \quad \text{and} \quad \hat{\mathcal{F}}_r(s) := \frac{\mathcal{F}_r(s)}{\mathcal{F}_r(-\rho)} = \frac{\mathcal{B}(-s, -r)}{\mathcal{B}(\rho, -r)}.$$

For parameters $a_1, \dots, a_N, b_1, \dots, b_R$ so that $a_i + b_j < 0$, the Heckman-Opdam measure on $r_1 \geq \dots \geq r_N > 0$ is given by the density $\mathcal{F}_r(a) \mathcal{F}_r(b) Z(a, b)^{-1}$ for the partition function

$$(4.2) \quad Z(a, b) := \prod_{i=1}^N \prod_{j=1}^R \frac{1}{-a_i - b_j} = \int_{r_1 \geq \dots \geq r_N > 0} \mathcal{F}_r(a) \mathcal{F}_r(b) dr,$$

where the second equality is a continuous version of the Cauchy-Littlewood identity; we refer the reader to [BG15a] for a generalization to general β random matrices.

Theorem 4.6. Let $X_{\alpha, R, N}$ be a random matrix from the Jacobi ensemble with parameters α and R . If $R/N = \hat{R} + O(N^{-1})$ and $\alpha/N = \hat{\alpha} + O(N^{-1})$ as $N \rightarrow \infty$ for some $\hat{R} \geq 1$ and $\hat{\alpha} > 0$, then the log-spectral measure of $X_{\alpha, R, N}$ is CLT-appropriate with

$$\Psi(u) = \log \frac{\hat{\alpha} + u}{\hat{\alpha} + \hat{R} + u} \quad \text{and} \quad \Lambda(u, w) = 0.$$

²While for $\beta = 2$ these functions are essentially the same as multivariate Bessel functions, for other values of β the difference is more significant.

Proof. We first compute the multivariate Bessel generating function for $X_{\alpha,R,N}$. It was shown in [BG15a, Theorem 2.8] that if $\lambda_i^N = e^{-r_i^N}$ are drawn from the Jacobi ensemble with parameters α and R , then r_i^N have the law of a Heckman-Opdam measure with $a_i = 1 - i$ and $b_j = 1 - j - \alpha$. Therefore, using (4.2) we see that

$$\mathbb{E}_{r^N} [\hat{\mathcal{F}}_{r^N}(s)] = \frac{Z(s, b)}{Z(-\rho, b)} = \prod_{i=1}^N \prod_{j=1}^R \frac{2 - i - j - \alpha}{s_i + 1 - j - \alpha} = \prod_{i=1}^N \frac{(1 - i - \alpha)_R}{(s_i - \alpha)_R},$$

where $(x)_n := x(x-1)\cdots(x-n+1)$. This implies that the defining integral of the multivariate Bessel generating function converges and equals

$$\phi_{X_{\alpha,R,N}}(s) = \mathbb{E}_{r^N} \left[\frac{\mathcal{B}(s, -r^N)}{\mathcal{B}(\rho, -r^N)} \right] = \mathbb{E}_{r^N} [\hat{\mathcal{F}}_{r^N}(-s)] = \prod_{i=1}^N \frac{(1 - i - \alpha)_R}{(-s_i - \alpha)_R}.$$

In particular, we see that

$$\log \phi_{X_{\alpha,R,N}}(s) = \sum_{i=1}^N \log \frac{(-N + i - \alpha)_R}{(-s_i - \alpha)_R},$$

where we see that for $\hat{\rho}_i = \frac{N-i}{N}$ by Stirling's approximation and the fact that $\alpha > 0$ we have

$$\begin{aligned} N^{-1} \log \frac{(-N + i - \alpha)_R}{(-rN - \alpha)_R} &= N^{-1} \log \frac{\Gamma(\alpha + N - i + R)\Gamma(\alpha + rN)}{\Gamma(\alpha + N - i)\Gamma(\alpha + rN + R)} \\ &= \tilde{\Psi}(r) - \tilde{\Psi}(\hat{\rho}_i) + O(N^{-1}) \end{aligned}$$

uniformly on a neighborhood of $[0, 1]$ for

$$\tilde{\Psi}(r) := (\hat{\alpha} + r) \log(\hat{\alpha} + r) - (\hat{\alpha} + \hat{R} + r) \log(\hat{\alpha} + \hat{R} + r).$$

This implies that

$$\frac{1}{N} \partial_{r_i} \log \phi_{X_{\alpha,R,N}}^I(rN) = \tilde{\Psi}'(r_i) + O(N^{-1})$$

uniformly in r_i and I . Since all mixed partials of $\phi_{X_{\alpha,R,N}}(s)$ vanish identically, this implies the desired. $\square$

Using the interpretation of the Jacobi ensemble as the squared singular values of submatrices of unitary matrices, we may take a limit to Wishart ensembles. Recall that the $L \times N$ complex Ginibre ensemble is the $L \times N$ random matrix of i.i.d. centered complex Gaussians ξ_{ij} with variance $\mathbb{E}[|\xi_{ij}|^2] = N^{-1}$ and that the $N \times N$ complex Wishart ensemble with parameter L is the corresponding random matrix $X_{L,N} := G_{L,N}^* G_{L,N}$.

Theorem 4.7. Let $G_{L,N}$ be a sequence of $L \times N$ complex Ginibre matrices, and let $X_{L,N} = G_{L,N}^* G_{L,N}$ be the corresponding sequence of complex Wishart matrices. If $L/N = \gamma + O(N^{-1})$ for some $\gamma > 1$, then the log-spectral measure of $X_{L,N}$ is CLT-appropriate with

$$\Psi(u) = \log(u + \gamma - 1) \quad \text{and} \quad \Lambda(u, w) = 0.$$

Proof. Set $\alpha = L - N + 1$. As $R \rightarrow \infty$ with all other dimensions fixed, we see by [PR04, Theorem 6.1] that a $L \times N$ submatrix of a Haar distributed element of U_{L+R} converges in distribution to a $L \times N$ complex Ginibre matrix scaled by $(L + R)^{-1/2} N^{1/2}$. Thus if $X_{\alpha,R,N}$ is drawn from the Jacobi ensemble, taking the limit as $R \rightarrow \infty$ yields a matrix $X_{L,N}$ from the corresponding Wishart ensemble. This implies that the

defining integral for $\phi_{X_{L,N}}(s)$ converges and equals

$$\begin{aligned}
\phi_{X_{L,N}}(s) &= \lim_{R \rightarrow \infty} \mathbb{E}_r \left[\frac{\mathcal{B}(s, \log(L+R) - \log(N) - r)}{\mathcal{B}(\rho, \log(L+R) - \log(N) - r)} \right] \\
&= \lim_{R \rightarrow \infty} \prod_{i=1}^N \left[N^{\rho_i - s_i} (L+R)^{s_i - \rho_i} \frac{(-i - L + N)_{L+R}}{(-s_i - L + N - 1)_{L+R}} \right] \\
&= \lim_{R \rightarrow \infty} \prod_{i=1}^N \left[N^{\rho_i - s_i} (L+R)^{s_i - \rho_i} \frac{(-L + i - 1)_{L+R}}{(-s_i - L + N - 1)_{L+R}} \right] \\
&= \lim_{R \rightarrow \infty} \prod_{i=1}^N \left[N^{\rho_i - s_i} (L+R)^{s_i - \rho_i} \frac{\Gamma(2L + R - i + 1) \Gamma(L - N + s_i + 1)}{\Gamma(L - i + 1) \Gamma(2L + R - N + s_i + 1)} \right] \\
&= \prod_{i=1}^N \frac{\Gamma(L - N + s_i + 1)}{\Gamma(L - i + 1)} N^{\rho_i - s_i}.
\end{aligned}$$

Applying Stirling's approximation and using the fact that $\gamma > 1$, we see that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \log \left[\frac{\Gamma(L - N + rN + 1)}{\Gamma(L - i + 1)} N^{\rho_i - rN} \right] = \tilde{\Psi}(r) - \tilde{\Psi}(\hat{\rho}_i) + O(N^{-1})$$

uniformly in r on a neighborhood of $[0, 1]$ for $\hat{\rho}_i = \frac{N-i}{N}$ and

$$\tilde{\Psi}(u) := (u + \gamma - 1) \log(u + \gamma - 1) - (u + \gamma - 1).$$

As before, this implies that

$$\frac{1}{N} \partial_{r_i} \log \phi_{X_{L,N}}^I(rN) = \tilde{\Psi}'(r_i) + O(N^{-1})$$

uniformly in r_i and I . Again, since all mixed partials of $\phi_{X_{L,N}}(s)$ vanish, this implies the desired. $\square$

4.3. Products of finitely many random matrices. We now apply Theorems 4.5, 4.6, and 4.7 to prove results on products of random matrices in several different settings. Suppose that $Y_N^1, \dots, Y_N^M$ are $N \times N$ random matrices which are right unitarily invariant, and let $X_N^i := (Y_N^i)^* Y_N^i$. Under different assumptions on the spectral measure of X_N^i , we study the spectral measure of

$$X_N := (Y_N^1 \cdots Y_N^M)^* (Y_N^1 \cdots Y_N^M)$$

where $N \rightarrow \infty$ and M is fixed. Let the eigenvalues of X_N be $\mu_1^N \geq \cdots \geq \mu_N^N > 0$, let their logarithms be $\lambda_i^N := \log \mu_i^N$, let

$$d\lambda^N := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^N}$$

denote the empirical log-spectral measure of X_N , and define the corresponding height function by

$$\mathcal{H}_N(t) := \#\{\lambda_i^N \leq t\}.$$

Finally, recall that the free product of two measures ρ_1 and ρ_2 was defined in [Voi85, Voi87] as the unique operation $(d\rho_1, d\rho_2) \mapsto d\rho_1 \boxtimes d\rho_2$ on probability measures which is compatible with multiplication of S -transforms, meaning that

$$S_{\rho_1 \boxtimes \rho_2}(z) = S_{\rho_1}(z) S_{\rho_2}(z).$$

Theorem 4.8. If the X_N^i have deterministic spectrum with support contained within fixed finite intervals and converging weakly to compactly supported measures $d\rho_i$, then as $N \rightarrow \infty$ with M fixed, the empirical log-spectral measure $d\lambda^N$ converges in probability in the sense of moments to the measure $d\rho$ whose pushforward $d\tilde{\rho}$ under the exponential map satisfies

$$d\tilde{\rho} = d\tilde{\rho}_1 \boxtimes \cdots \boxtimes d\tilde{\rho}_M.$$

Its centered moments $\{p_k(\lambda) - \mathbb{E}[p_k(\lambda)]\}_{k \in \mathbb{N}}$ converge in probability to a Gaussian vector with covariance

$$\text{Cov}\left(p_k(\lambda), p_l(\lambda)\right) = \oint \oint \left(\log(u/(u-1)) - \log S_{\tilde{\rho}}(u-1) \right)^k \left(\log(w/(w-1)) - \log S_{\tilde{\rho}}(w-1) \right)^l \left(\sum_{i=1}^M \frac{1}{M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(u-1))M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(w-1))(M_{\tilde{\rho}_i}^{-1}(u-1) - M_{\tilde{\rho}_i}^{-1}(w-1))^2} - \frac{M-1}{(u-w)^2} \right).$$

Proof. By Theorem 4.5, the log-spectral measure of each X_N^i is CLT-appropriate, so by Lemma 4.4 the log-spectral measure of X_N is CLT-appropriate with

$$\Psi(u) = -\log\left(S_{\tilde{\rho}_1}(u-1) \cdots S_{\tilde{\rho}_M}(u-1)\right)$$

$$\Lambda(u, w) = \sum_{i=1}^M \left[\frac{1}{M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(u-1))M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(w-1))(M_{\tilde{\rho}_i}^{-1}(u-1) - M_{\tilde{\rho}_i}^{-1}(w-1))^2} - \frac{1}{(u-w)^2} \right].$$

Define the measure $d\tilde{\rho}$ as the free product

$$d\tilde{\rho} := d\tilde{\rho}_1 \boxtimes \cdots \boxtimes d\tilde{\rho}_M$$

so that

$$S_{\tilde{\rho}}(u) = S_{\tilde{\rho}_1}(u) \cdots S_{\tilde{\rho}_M}(u).$$

By Theorem 2.5 and Lemma 2.7, the log-spectral measure of X_N converges in probability in the sense of moments to a measure with k^{th} moment

$$\begin{aligned} \mathfrak{p}_k &= \frac{1}{k+1} \oint \left(\log(u/(u-1)) - \log S_{\tilde{\rho}}(u-1) \right)^{k+1} \frac{du}{2\pi\mathbf{i}} \\ &= \frac{1}{k+1} \oint [-\log M_{\tilde{\rho}}^{-1}(u-1)]^{k+1} \frac{du}{2\pi\mathbf{i}} \\ &= \frac{1}{k+1} \int \oint [-\log z]^{k+1} \frac{e^{-s}}{(e^{-s}-z)^2} \frac{dz}{2\pi\mathbf{i}} d\rho(s) \\ &= \int s^k d\rho(s), \end{aligned}$$

where we make the change of variable $z = M_{\tilde{\rho}}^{-1}(u-1)$ so that $u = M_{\tilde{\rho}}(z) + 1$, $du = M'_{\tilde{\rho}}(z)dz$, and the z -contour is clockwise around $\{z^{-1} \mid z \in \text{supp } d\tilde{\rho}\}$. Because $d\rho(s)$ is compactly supported, convergence of moments implies convergence of random measures, as desired. The fluctuations of moments now follow from Theorem 2.6, Lemma 2.7, and the fact that the log-spectral measure is CLT-appropriate. $\square$

If $d\rho(s) = p(s)ds$ is absolutely continuous with α -Hölder continuous density $p(s)$ with respect to Lebesgue measure, meaning that $|p(x) - p(y)| < C|x - y|^\alpha$ for $|x - y| < \delta$ for some $\alpha > 0$, $\delta > 0$ and $C > 0$, we may define the Cauchy principal value integral

$$M_{\tilde{\rho}}(e^{-t}) := \text{p.v.} \int \frac{e^{s-t}}{1 - e^{s-t}} d\rho(s),$$

which satisfies

$$(4.3) \quad \lim_{\varepsilon \rightarrow 0^\pm} M_{\tilde{\rho}}(e^{-t+i\varepsilon}) = M_{\tilde{\rho}}(e^{-t}) \pm i\pi p(t)$$

by virtue of the relation

$$M_{\tilde{\rho}}(e^{-t}) = \int \frac{1}{e^t - r} p(\log r) dr$$

between $M_{\tilde{\rho}}(e^{-t})$ and the Stieltjes transform of the measure with density $p(\log r)$.

Corollary 4.9. Suppose that the measure $d\rho$ is absolutely continuous with respect to Lebesgue measure with α -Hölder continuous density $d\rho = p(s)ds$. The centered height function $\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]$ converges in

the sense of moments to the Gaussian random field on $\mathbb{R}$ with covariance

$$K(t, s) = -\frac{1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right| \\ - \frac{1}{2\pi^2} \sum_{i=1}^M \left[\log \left| \frac{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi p(s))}{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) - \mathbf{i}\pi p(s))} \right| - \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right| \right].$$

Proof. Choose an interval $I = [t_1, t_2]$ so that $\text{supp } d\lambda^N \subset I$. By integration by parts we see that

$$\int_I \mathcal{H}_N(t) t^k dt = \frac{\mathcal{H}_N(t_2)}{k+1} t_2^{k+1} - \frac{\mathcal{H}_N(t_1)}{k+1} t_1^{k+1} - \int_I \mathcal{H}'_N(t) \frac{t^{k+1}}{k+1} dt \\ = \frac{\mathcal{H}_N(t_2)}{k+1} t_2^{k+1} - \frac{\mathcal{H}_N(t_1)}{k+1} t_1^{k+1} - \frac{1}{k+1} p_{k+1}(\lambda^N),$$

which means that

$$H_k := \int_I (\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) t^k dt = -\frac{1}{k+1} (p_{k+1}(\lambda^N) - \mathbb{E}[p_{k+1}(\lambda^N)]).$$

By Theorem 4.8, we find that H_k are asymptotically Gaussian with covariance $\text{Cov}(H_k, H_l)$ given by

$$\frac{1}{(k+1)(l+1)} \mathfrak{Cov}_{k+1, l+1} = \frac{1}{(k+1)(l+1)} \oint \oint \left(\log(u/(u-1)) - \log S_{\tilde{\rho}}(u-1) \right)^{l+1} \\ \left(\log(w/(w-1)) - \log S_{\tilde{\rho}}(w-1) \right)^{k+1} G^{(1,1)}(u, w) \frac{du}{2\pi \mathbf{i}} \frac{dw}{2\pi \mathbf{i}} \\ = \frac{1}{(k+1)(l+1)} \oint \oint [-\log M_{\tilde{\rho}}^{-1}(u-1)]^{l+1} [-\log M_{\tilde{\rho}}^{-1}(w-1)]^{k+1} G^{(1,1)}(u, w) \frac{du}{2\pi \mathbf{i}} \frac{dw}{2\pi \mathbf{i}} \\ = \frac{1}{(k+1)(l+1)} \oint \oint z^{l+1} v^{k+1} \partial_z \partial_v G(M_{\tilde{\rho}}(e^{-z}) + 1, M_{\tilde{\rho}}(e^{-v}) + 1) \frac{dz}{2\pi \mathbf{i}} \frac{dv}{2\pi \mathbf{i}} \\ = \oint \oint z^l v^k H(z, v) \frac{dz}{2\pi \mathbf{i}} \frac{dv}{2\pi \mathbf{i}} \\ = \frac{1}{(2\pi \mathbf{i})^2} \lim_{\varepsilon \rightarrow 0} \int_{\text{supp } d\rho} \int_{\text{supp } d\rho} z^l v^k [H(z + \mathbf{i}\varepsilon, v + \mathbf{i}\varepsilon) - H(z - \mathbf{i}\varepsilon, v + \mathbf{i}\varepsilon) \\ - H(z + \mathbf{i}\varepsilon, v - \mathbf{i}\varepsilon) + H(z - \mathbf{i}\varepsilon, v - \mathbf{i}\varepsilon)] dz dv$$

for

$$G(u, w) = \sum_{i=1}^M \left[\log(M_{\tilde{\rho}_i}^{-1}(u-1) - M_{\tilde{\rho}_i}^{-1}(w-1)) - \log(u-w) \right] + \log(u-w) \\ G^{(1,1)}(u, w) = \left(\sum_{i=1}^M \frac{1}{M_{\tilde{\rho}_i}'(M_{\tilde{\rho}_i}^{-1}(u-1)) M_{\tilde{\rho}_i}'(M_{\tilde{\rho}_i}^{-1}(w-1)) (M_{\tilde{\rho}_i}^{-1}(u-1) - M_{\tilde{\rho}_i}^{-1}(w-1))^2} - \frac{M-1}{(u-w)^2} \right) \\ H(z, v) = G(M_{\tilde{\rho}}(e^{-z}) + 1, M_{\tilde{\rho}}(e^{-v}) + 1),$$

which means that the covariance of the relevant Gaussian field is given by

$$K(t, s) := \lim_{\varepsilon \rightarrow 0^+} -\frac{1}{4\pi^2} [H(t + \mathbf{i}\varepsilon, s + \mathbf{i}\varepsilon) - H(t + \mathbf{i}\varepsilon, s - \mathbf{i}\varepsilon) - H(t - \mathbf{i}\varepsilon, s + \mathbf{i}\varepsilon) + H(t - \mathbf{i}\varepsilon, s - \mathbf{i}\varepsilon)] \\ = \lim_{\varepsilon \rightarrow 0^+} -\frac{1}{2\pi^2} \Re [H(t + \mathbf{i}\varepsilon, s + \mathbf{i}\varepsilon) - H(t + \mathbf{i}\varepsilon, s - \mathbf{i}\varepsilon)].$$

By (4.3), we find that

$$\lim_{\varepsilon \rightarrow 0^+} \log \left| \frac{M_{\tilde{\rho}}(e^{-t+\mathbf{i}\varepsilon}) - M_{\tilde{\rho}}(e^{-s+\mathbf{i}\varepsilon})}{M_{\tilde{\rho}}(e^{-t+\mathbf{i}\varepsilon}) - M_{\tilde{\rho}}(e^{-s-\mathbf{i}\varepsilon})} \right| = \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right|.$$

Similarly, we see that

$$\lim_{\varepsilon \rightarrow 0^+} \log \left| \frac{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t+\mathbf{i}\varepsilon})) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s+\mathbf{i}\varepsilon}))}{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t+\mathbf{i}\varepsilon})) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s-\mathbf{i}\varepsilon}))} \right| = \log \left| \frac{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi p(s))}{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) - \mathbf{i}\pi p(s))} \right|.$$

Putting these together, we conclude that

$$K(t, s) = -\frac{1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right| \\ - \frac{1}{2\pi^2} \sum_{i=1}^M \left[\log \left| \frac{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi p(s))}{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) - \mathbf{i}\pi p(s))} \right| - \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right| \right]. \quad \square$$

Remark. The function $\omega_i := M_{\tilde{\rho}_i}^{-1} \circ M_{\tilde{\rho}}$ on $\mathbb{C} - [0, \infty)$ which appears in the expression

$$M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) \pm \mathbf{i}\pi p(t)) = \lim_{\varepsilon \rightarrow 0^\pm} M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t+\mathbf{i}\varepsilon}))$$

in Theorem 4.8 is the subordination function corresponding to multiplicative convolution of measures as defined in [Voi93, Bia98, BB07]. We relate this function to $\text{supp } d\tilde{\rho}$ in the following Lemma 4.10.

Lemma 4.10. If the measures $d\tilde{\rho}$ and $d\tilde{\rho}_i$ are absolutely continuous with respect to Lebesgue measure with α -Hölder continuous densities, then $\lim_{\varepsilon \rightarrow 0^+} \omega_i(t + \mathbf{i}\varepsilon)$ is real for some $t \in [0, \infty)$ if and only if $t^{-1} \notin \text{supp } d\tilde{\rho}$.

Proof. If $t^{-1} \notin \text{supp } d\tilde{\rho}$, the statement follows from [BBCF17, Lemma 3.2]. Now, suppose that $\lim_{\varepsilon \rightarrow 0^+} \omega_i(t + \mathbf{i}\varepsilon)$ is real; if $t = 0$, we are done since $d\tilde{\rho}$ is compactly supported. Otherwise, suppose for the sake of contradiction that $t^{-1} \in \text{supp } d\tilde{\rho}$. Define

$$d\tilde{\rho}_i^* = d\tilde{\rho}_1 \boxtimes \cdots \boxtimes \widehat{d\tilde{\rho}_i} \boxtimes \cdots \boxtimes d\tilde{\rho}_M$$

so that

$$d\tilde{\rho} = d\tilde{\rho}_i \boxtimes d\tilde{\rho}_i^*.$$

Letting ω_i^* denote the subordination function for $d\tilde{\rho}_i^*$, we get from [BBCF17, Equation 3.7] that for $z \in \mathbb{C} \setminus [0, \infty)$ we have

$$(4.4) \quad \omega_i(z)\omega_i^*(z) = z\eta_{\tilde{\rho}}(z) = z\eta_{\tilde{\rho}_i}(\omega_i^*(z)) = z\eta_{\tilde{\rho}_i^*}(\omega_i(z))$$

for

$$\eta_{\tilde{\rho}}(z) := \frac{M_{\tilde{\rho}}(z)}{1 + M_{\tilde{\rho}}(z)}.$$

Equating the second and fourth terms in (4.4), we find that $M_{\tilde{\rho}}(z) = M_{\tilde{\rho}_i}(\omega_i(z))$ for $z \in \mathbb{C} \setminus [0, \infty)$. In particular, this means that

$$\lim_{\varepsilon \rightarrow 0^+} [M_{\tilde{\rho}}(t + \mathbf{i}\varepsilon) - M_{\tilde{\rho}}(t - \mathbf{i}\varepsilon)] = \lim_{\varepsilon \rightarrow 0^+} [M_{\tilde{\rho}_i}(\omega_i(t + \mathbf{i}\varepsilon)) - M_{\tilde{\rho}_i}(\omega_i(t - \mathbf{i}\varepsilon))].$$

Since $t^{-1} \in \text{supp } d\tilde{\rho}$ and $d\tilde{\rho}$ has a α -Hölder continuous density, this expression is nonzero by (4.3), hence we must have $\lim_{\varepsilon \rightarrow 0^+} \omega_i(t + \mathbf{i}\varepsilon) \in [0, \infty)$. In this case, we see that

$$\lim_{\varepsilon \rightarrow 0^+} \arg \omega_i^*(t + \mathbf{i}\varepsilon) = \lim_{\varepsilon \rightarrow 0^+} \arg \frac{(t + \mathbf{i}\varepsilon)\eta_{\tilde{\rho}_i}(\omega_i^*(t + \mathbf{i}\varepsilon))}{\omega_i(t + \mathbf{i}\varepsilon)} = \lim_{\varepsilon \rightarrow 0^+} \arg \eta_{\tilde{\rho}_i}(\omega_i^*(t + \mathbf{i}\varepsilon)).$$

Now, if $w = \lim_{\varepsilon \rightarrow 0^+} \omega_i^*(t + \mathbf{i}\varepsilon) \notin [0, \infty)$, then $\arg \eta_{\tilde{\rho}_i}$ is continuous at w and hence

$$\arg w = \arg \eta_{\tilde{\rho}_i}(w),$$

which implies by property (a) of $\eta_{\tilde{\rho}_i}(z)$ in [BBCF17, Section 3.2] that $\tilde{\rho}_i$ is a point mass, which is excluded by the given. We conclude that $\lim_{\varepsilon \rightarrow 0^+} \omega_i^*(t + \mathbf{i}\varepsilon) \in [0, \infty)$. Substituting $z = t + \mathbf{i}\varepsilon$ into the first equality in (4.4) and taking the limit as $\varepsilon \rightarrow 0^+$, we find that

$$\lim_{\varepsilon \rightarrow 0^+} \eta_{\tilde{\rho}}(t + \mathbf{i}\varepsilon) \in \mathbb{R},$$

which implies by virtue of the α -Hölder continuous density of $d\tilde{\rho}$ that $t^{-1} \notin \text{supp } d\tilde{\rho}$, as desired. $\square$

Corollary 4.11. Suppose that the density $p(s)$ for $d\rho(s)$ is $C^{1+\alpha}$ for some $\alpha > 0$ on the interior of $\text{supp } d\rho$ and no $d\tilde{\rho}_i$ contains a point mass. Then for s in the interior of $\text{supp } d\rho$, the covariance of Corollary 4.9 satisfies

$$K(t, s) = -\frac{1}{2\pi^2} \log |t - s| - \frac{1}{2\pi^2} \log \left| \frac{\partial_t [M_{\tilde{\rho}}(e^{-t})] + \mathbf{i}\pi p'(t)}{2\pi p(t)} \right| \\ + \frac{1}{2\pi^2} \sum_{i=1}^M \log \left| \frac{(M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) - \mathbf{i}\pi p(t))) M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)))}{2\pi p(t)} \right| + o(t - s)$$

as $(t - s) \rightarrow 0$, meaning that it matches that of a log-correlated Gaussian field along the diagonal.

Proof. By exchanging the derivative and Cauchy principal value integral, we see that

$$\partial_t M_{\tilde{\rho}}(e^{-t}) = \text{p.v.} \int \frac{1}{1 - e^x} p'(x + t) dx$$

exists if $p(s)$ is $C^{1+\alpha}$ near $s = t$. This implies that for s in the interior of $\text{supp } d\rho$, as $t \rightarrow s$ we have

$$(4.5) \quad \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right| = \log \left| \frac{(t - s)[\partial_t [M_{\tilde{\rho}}(e^{-t})] + \mathbf{i}\pi p'(t)]}{(t - s)[\partial_t [M_{\tilde{\rho}}(e^{-t})] + \mathbf{i}\pi p'(t)] - 2\mathbf{i}\pi p(t)} \right| + o(t - s) \\ = \log |t - s| + \log \left| \frac{\partial_t [M_{\tilde{\rho}}(e^{-t})] + \mathbf{i}\pi p'(t)}{2\pi p(t)} \right| + o(t - s).$$

Noting that $\partial_t M_{\tilde{\rho}_i}^{-1}(t) = \frac{1}{M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(t))}$, we find therefore that as $t \rightarrow s$ we have

$$M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi p(s)) \\ = (t - s)[\partial_t M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p'(t)] \frac{1}{M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)))} + o(t - s) \\ M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) - \mathbf{i}\pi p(s)) \\ = M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) - \mathbf{i}\pi p(t)) + o(t - s).$$

Putting these together, if $M = 1$, we see that $K(t, s) = 0$; if $M > 1$, we see that as $t \rightarrow s$, we have that

$$(4.6) \quad \log \left| \frac{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi p(s))}{M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-s}) - \mathbf{i}\pi p(s))} \right| - \log \left| \frac{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) - p(s))}{M_{\tilde{\rho}}(e^{-t}) - M_{\tilde{\rho}}(e^{-s}) + \mathbf{i}\pi(p(t) + p(s))} \right| \\ = \log \left| \frac{(M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) - \mathbf{i}\pi p(t))) M'_{\tilde{\rho}_i}(M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)))}{2\pi p(t)} \right| + o(t - s),$$

where since $e^t \in \text{supp } d\tilde{\rho}$, by Lemma 4.10 we have for $M > 1$ that

$$\lim_{\varepsilon \rightarrow 0^+} [\omega_i(t + \mathbf{i}\varepsilon) - \omega_i(t - \mathbf{i}\varepsilon)] = M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) + \mathbf{i}\pi p(t)) - M_{\tilde{\rho}_i}^{-1}(M_{\tilde{\rho}}(e^{-t}) - \mathbf{i}\pi p(t)) \neq 0.$$

Adding (4.5) and the summation of (4.6) over all i yields the claim. $\square$

Remark. The additive analogue of Theorem 4.8 and Corollary 4.9 for two matrices was studied in [MSS07], [CMSS07], [MS17, Chapter 5], [PS11, Chapter 10], and [BG18a, Proposition 9.8]. They considered unitarily invariant Hermitian random matrices $A_N^1, \dots, A_N^M$ such that the empirical spectral measures $d\rho_{A^i}$ of A_N^i converge weakly to deterministic measures $d\rho_{A^i}$. Define the R -transform and Cauchy transform of a measure ρ by

$$R_\rho(z) := G_\rho^{-1}(z) - z^{-1} \quad \text{and} \quad G_\rho(z) := \int \frac{1}{z - s} d\rho(s)$$

Then the empirical spectral measure of $X_M = \sum_{i=1}^M A_N^i$ concentrates on a measure $d\rho_{X_M} = d\rho_{A^1} \boxplus \dots \boxplus d\rho_{A^M}$ for which

$$R_{\rho_{X_M}}(z) = \sum_{i=1}^M R_{\rho_{A^i}}(z).$$

In [PS11, Chapter 10], it was shown further that the fluctuations of the spectrum λ^N of C_N satisfy

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}(p_k(\lambda^N), p_l(\lambda^N)) &= \oint \oint z^k w^l \partial_z \partial_w \left[\sum_{i=1}^M \log \left(G_{\rho_{A^i}}^{-1}(G_{\rho_{X_M}}(z)) - G_{\rho_{A^i}}^{-1}(G_{\rho_{X_M}}(w)) \right) \right. \\ &\quad \left. - \log(z-w) - (M-1) \log \left(G_{\rho_{X_M}}(z)^{-1} - G_{\rho_{X_M}}(w)^{-1} \right) \right] \frac{dz}{2\pi i} \frac{dw}{2\pi i}, \end{aligned}$$

where both contours enclose only the poles at ∞ . Applying the same reasoning as in the proof of Theorem 4.8, this implies that the height function

$$\mathcal{H}_N(t) := \#\{\lambda_i^N \leq t\}$$

has moments

$$H_k := \int (\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) t^k dt$$

with limiting covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}(H_k, H_l) &= \frac{1}{(k+1)(l+1)} \oint \oint z^{k+1} w^{l+1} \partial_z \partial_w \left[\sum_{i=1}^M \log \left(G_{\rho_{A^i}}^{-1}(G_{\rho_{X_M}}(z)) - G_{\rho_{A^i}}^{-1}(G_{\rho_{X_M}}(w)) \right) \right. \\ &\quad \left. - \log(z-w) - (M-1) \log \left(G_{\rho_{X_M}}(z)^{-1} - G_{\rho_{X_M}}(w)^{-1} \right) \right] \frac{dz}{2\pi i} \frac{dw}{2\pi i} \\ &= \oint \oint z^k w^l \left[\sum_{i=1}^M \log \left(G_{\rho_{A^i}}^{-1}(G_{\rho_{X_M}}(z)) - G_{\rho_{A^i}}^{-1}(G_{\rho_{X_M}}(w)) \right) \right. \\ &\quad \left. - \log(z-w) - (M-1) \log \left(G_{\rho_{X_M}}(z)^{-1} - G_{\rho_{X_M}}(w)^{-1} \right) \right] \frac{dz}{2\pi i} \frac{dw}{2\pi i} \\ &= -\frac{1}{2\pi^2} \int_I \int_I z^k w^l \sum_{i=1}^M \left[\log \left| \frac{G_{\rho_{A^i}}^{-1}(H_{X_M}(z)) - G_{\rho_{A^i}}^{-1}(H_{X_M}(w))}{G_{\rho_{A^i}}^{-1}(H_{X_M}(z)) - \overline{G_{\rho_{A^i}}^{-1}(H_{X_M}(w))}} \right| - \log \left| \frac{H_{X_M}(z)^{-1} - H_{X_M}(w)^{-1}}{H_{X_M}(z)^{-1} - \overline{H_{X_M}(w)^{-1}}} \right| \right] dz dw \\ &\quad - \frac{1}{2\pi^2} \int_I \int_I z^k w^l \log \left| \frac{H_{X_M}(z)^{-1} - H_{X_M}(w)^{-1}}{H_{X_M}(z)^{-1} - \overline{H_{X_M}(w)^{-1}}} \right| dz dw, \end{aligned}$$

where $I = \text{supp } d\rho_{X_M}$ and

$$H_{X_M}(z) := \lim_{\varepsilon \rightarrow 0^+} G_{\rho_{X_M}}(z + i\varepsilon).$$

This means that $\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]$ converges to the Gaussian field on $\text{supp } d\rho_{X_M}$ with covariance

$$(4.7) \quad K(t, s) := -\frac{1}{2\pi^2} \sum_{i=1}^M \left[\log \left| \frac{G_{\rho_{A^i}}^{-1}(H_{X_M}(t)) - G_{\rho_{A^i}}^{-1}(H_{X_M}(s))}{G_{\rho_{A^i}}^{-1}(H_{X_M}(t)) - \overline{G_{\rho_{A^i}}^{-1}(H_{X_M}(s))}} \right| - \log \left| \frac{H_{X_M}(t)^{-1} - H_{X_M}(s)^{-1}}{H_{X_M}(t)^{-1} - \overline{H_{X_M}(s)^{-1}}} \right| \right] \\ - \frac{1}{2\pi^2} \log \left| \frac{H_{X_M}(t)^{-1} - H_{X_M}(s)^{-1}}{H_{X_M}(t)^{-1} - \overline{H_{X_M}(s)^{-1}}} \right|.$$

Theorem 4.12. If the X_N^i have spectrum which is a Jacobi ensemble with parameters α and R with $R/N = \hat{R} + O(N^{-1})$ and $\alpha/N = \hat{\alpha} + O(N^{-1})$ for $\hat{R} \geq 1$ and $\hat{\alpha} > 0$, then as $N \rightarrow \infty$ the empirical log-spectral measure $d\lambda^N$ of X_N converges in probability in the sense of moments to the measure $d\rho$ whose pushforward $d\tilde{\rho}$ under the exponential map has S -transform

$$S_{\tilde{\rho}}(u) = \left(\frac{\hat{\alpha} + \hat{R} + u + 1}{\hat{\alpha} + u + 1} \right)^M.$$

For the height function defined by

$$\mathcal{H}_N(t) := \#\{e^{\lambda_i^N} \leq t\},$$

the centered height function $\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]$ converges in the sense of moments to the Gaussian random field on $[0, 1]$ with covariance

$$K(t, s) = -\frac{1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + i\pi(p(\log t) - p(\log s))}{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + i\pi(p(\log t) + p(\log s))} \right|,$$

where $p(s)$ is the density of $d\tilde{\rho}$ and $M_{\tilde{\rho}}(t)$ is defined on $[0, 1]$ in the Cauchy principal value sense by

$$M_{\tilde{\rho}}(t) := \text{p.v.} \int \frac{e^{st}}{1 - e^{st}} d\rho(s) \quad \text{and} \quad d\rho(s) = p(s)ds.$$

Proof. By Theorem 4.6, the log-spectral measure of each X_N^i is CLT-appropriate, so by Lemma 4.4 the log-spectral measure $d\lambda^N$ of X_N is CLT-appropriate with

$$\Psi(u) = M \log \frac{\hat{\alpha} + u}{\hat{\alpha} + \hat{R} + u} = -\log S_{\tilde{\rho}}(u - 1) \quad \text{and} \quad \Lambda(u, w) = 0.$$

By Theorem 2.5 and Lemma 2.7, the log-spectral measure of X_N converges in probability in the sense of moments to a measure with k^{th} moment

$$\mathfrak{p}_k = \frac{1}{k+1} \oint \left(\log(u/(u-1)) - \log S_{\tilde{\rho}}(u-1) \right)^{k+1} \frac{du}{2\pi i} = \int s^k d\rho(s)$$

as in the proof of Theorem 4.8. Because $d\rho(s)$ is compactly supported, convergence of moments implies convergence of random measures, as desired.

For the central limit theorem, by integration by parts we see that

$$\begin{aligned} \int_0^1 \mathcal{H}_N(t) t^k dt &= \int_{-\infty}^0 \mathcal{H}_N(e^s) e^{s(k+1)} ds \\ &= \frac{N}{k+1} - \frac{1}{k+1} \int_{-\infty}^0 \mathcal{H}'_N(e^s) e^{s(k+1)} ds \\ &= \frac{N}{k+1} - \frac{p_{k+1}(e^{\lambda^N})}{k+1}, \end{aligned}$$

which means that

$$H_k := \int_0^1 (\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) t^k dt = -\frac{1}{k+1} \left(p_{k+1}(e^{\lambda^N}) - \mathbb{E}[p_{k+1}(e^{\lambda^N})] \right).$$

By Theorem 2.6 and Lemma 2.7, we see that

$$\begin{aligned} \text{Cov}(H_k, H_l) &= \frac{1}{(k+1)(l+1)} \sum_{m,n=0}^{\infty} \text{Cov} \left(\frac{p_n(\lambda^N)(k+1)^n}{n!}, \frac{p_m(\lambda^N)(l+1)^m}{m!} \right) \\ &= \frac{1}{(k+1)(l+1)} \sum_{m,n=0}^{\infty} \frac{(k+1)^n (l+1)^m}{n! m!} \oint \oint \left(\log(u/(u-1)) - \log S_{\tilde{\rho}}(u-1) \right)^n \\ &\quad \left(\log(w/(w-1)) - \log S_{\tilde{\rho}}(w-1) \right)^m \frac{1}{(u-w)^2} \frac{du}{2\pi i} \frac{dw}{2\pi i} \\ &= \frac{1}{(k+1)(l+1)} \oint \oint M_{\tilde{\rho}}^{-1}(u-1)^{-k-1} M_{\tilde{\rho}}^{-1}(w-1)^{-l-1} \frac{1}{(u-w)^2} \frac{du}{2\pi i} \frac{dw}{2\pi i} \\ &= \frac{1}{(k+1)(l+1)} \oint \oint z^{k+1} v^{l+1} \frac{M'_{\tilde{\rho}}(z^{-1}) M'_{\tilde{\rho}}(v^{-1}) z^{-2} v^{-2}}{(M_{\tilde{\rho}}(z^{-1}) - M_{\tilde{\rho}}(v^{-1}))^2} \frac{dz}{2\pi i} \frac{dv}{2\pi i} \\ &= \oint \oint z^k v^l \log \left(M_{\tilde{\rho}}(z^{-1}) - M_{\tilde{\rho}}(v^{-1}) \right) \frac{dz}{2\pi i} \frac{dv}{2\pi i}. \end{aligned}$$

Applying the same argument as in the proof of Corollary 4.9, we find that the covariance of the relevant Gaussian field is given by

$$K(t, s) = -\frac{1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + i\pi(p(\log t) - p(\log s))}{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + i\pi(p(\log t) + p(\log s))} \right|. \quad \square$$

Remark. When $M = 1$, our proof of Theorem 4.12 recovers the central limit theorem for global fluctuations of the Jacobi ensemble. There are now several different proofs of this result (c.f. the discussion before [BG15a, Proposition 1.3]). We now match our covariance for this case with that of [BG15a] for $\beta = 2$. More specifically, we may compute

$$S_{\tilde{\rho}}(u) = \frac{\hat{\alpha} + \hat{R} + u + 1}{\hat{\alpha} + u + 1},$$

which implies that our computation of $\text{Cov}(H_k, H_l)$ coincides with the expression in [BG15a, Theorem 4.1] after changing variables to $u \mapsto -u$ and $w \mapsto -w$. Furthermore, we see that

$$M_{\tilde{\rho}}^{-1}(-u-1)^{-1} = \frac{u}{u+1} \frac{u-\hat{\alpha}}{u-\hat{\alpha}-\hat{R}},$$

which means that $-1 - M_{\tilde{\rho}}(u^{-1})$ coincides with $\Omega(u; 1)$ for $\hat{M} = \hat{R}$ in [BG15a, Definition 4.11]. If we deform the contours differently in the last step of Theorem 4.12 and make the change of variables $u \mapsto -u$ and $w \mapsto -w$, our result coincides with [BG15a, Theorem 4.13].

Theorem 4.13. If $X_N^i = (G_N^i)^* G_N^i$ with G_N^i a Ginibre ensemble with parameters L and N with $L/N = \gamma + O(N^{-1})$, then as $N \rightarrow \infty$, the empirical log-spectral measure $d\lambda^N$ of X_N converges in probability in the sense of moments to the measure $d\rho$ whose pushforward under the exponential map has S -transform

$$S_{\tilde{\rho}}(u) = \frac{1}{(u + \gamma)^M}.$$

For the height function defined by

$$\mathcal{H}_N(t) := \#\{e^{\lambda_i^N} \leq t\},$$

the centered height function $\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]$ converges in the sense of moments to the Gaussian random field on $[0, \infty)$ with covariance

$$K(t, s) = -\frac{1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + \mathbf{i}\pi(p(\log t) - p(\log s))}{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + \mathbf{i}\pi(p(\log t) + p(\log s))} \right|,$$

where $p(s)$ is the density of $d\tilde{\rho}$ and $M_{\tilde{\rho}}(t)$ is defined on $[0, \infty)$ by the Cauchy principal value integral

$$M_{\tilde{\rho}}(t) = \text{p.v.} \int \frac{e^{st}}{1 - e^{st}} d\rho(t) \quad \text{and} \quad d\rho(s) = p(s)ds.$$

Proof. By Theorem 4.7, the log-spectral measure of each X_N^i is CLT-appropriate, so by Lemma 4.4 the measure $d\lambda^N$ is CLT-appropriate with

$$\Psi(u) = M \log(u + \gamma - 1) = -\log S_{\tilde{\rho}}(u - 1) \quad \text{and} \quad \Lambda(u, w) = 0.$$

As in Theorem 4.12, by Theorem 2.5, the log-spectral measure of X_N converges in probability in the sense of moments to a measure with k^{th} moment which matches those of $d\rho$. Because $d\rho(s)$ is compactly supported, convergence of moments implies convergence of random measures, as desired.

For the central limit theorem, similarly to Theorem 4.6, by integration by parts we see that

$$\int_0^\infty \mathcal{H}_N(t) t^k dt = \frac{N}{k+1} - \frac{p_{k+1}(e^{\lambda^N})}{k+1},$$

which means that

$$H_k := \int_0^\infty (\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) t^k dt = -\frac{1}{k+1} (p_{k+1}(e^{\lambda^N}) - \mathbb{E}[p_{k+1}(e^{\lambda^N})]).$$

By Theorem 2.6 and Lemma 2.7, we see that

$$\begin{aligned} \text{Cov}(H_k, H_l) &= \frac{1}{(k+1)(l+1)} \sum_{m,n=0}^\infty \frac{(k+1)^n (l+1)^m}{n!m!} \text{Cov}(p_n(\lambda^N), p_m(\lambda^N)) \\ &= \frac{1}{(k+1)(l+1)} \sum_{m,n=0}^\infty \frac{(k+1)^n (l+1)^m}{n!m!} \oint \oint \left(\log(u/(u-1)) - \log S_{\tilde{\rho}}(u-1) \right)^n \\ &\quad \left(\log(w/(w-1)) - \log S_{\tilde{\rho}}(w-1) \right)^m \frac{1}{(u-w)^2} \frac{du}{2\pi\mathbf{i}} \frac{dw}{2\pi\mathbf{i}} \\ &= \oint \oint z^k v^l \log \left(M_{\tilde{\rho}}(z^{-1}) - M_{\tilde{\rho}}(v^{-1}) \right) \frac{dz}{2\pi\mathbf{i}} \frac{dv}{2\pi\mathbf{i}}, \end{aligned}$$

where we apply the exact same arguments as in the proof of Theorem 4.6. We therefore obtain that the covariance of the relevant Gaussian field is given by

$$K(t, s) = -\frac{1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + \mathbf{i}\pi(p(\log t) - p(\log s))}{M_{\tilde{\rho}}(t^{-1}) - M_{\tilde{\rho}}(s^{-1}) + \mathbf{i}\pi(p(\log t) + p(\log s))} \right|. \quad \square$$

Remark. When $M = 1$, $d\tilde{\rho}$ is the spectral measure of the Marchenko-Pastur distribution, as expected. Our computation of the fluctuations of the moments of the spectral measure agrees with the single-level case of [DP18, Proposition 1.2] after an integration by parts and deformation of contours.

4.4. Lyapunov exponents for matrices with fixed spectrum. We now apply our theorems to products of infinitely many random matrices. Suppose that $Y_N^1, \dots, Y_N^M$ are $N \times N$ random matrices which are right unitarily invariant, and let $X_N^i := (Y_N^i)^* Y_N^i$. As before, we study the spectral measure of

$$X_N := (Y_N^1 \cdots Y_N^M)^* (Y_N^1 \cdots Y_N^M),$$

where we now make the assumption that $M, N \rightarrow \infty$ simultaneously. In this setting, we expect the eigenvalues to grow exponentially in M , so we study the Lyapunov exponents

$$\lambda_i^N := \frac{1}{M} \log \mu_i^N,$$

where $\mu_1^N \geq \dots \geq \mu_N^N$ are the eigenvalues of X_N . Define their empirical measure by

$$d\lambda^N := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^N}$$

In this setting, we may identify the limiting measure as the sum of a deterministic measure and an explicit Gaussian process under rescaling. For $t \in \mathbb{R}$, define the height function

$$\mathcal{H}_N(t) := \#\{\lambda_i^N \leq t\}.$$

Theorem 4.14. If the X_N^i have deterministic spectrum satisfying Assumption 3.3 for a compactly supported non-atomic measure $d\rho$, then as $M, N \rightarrow \infty$, the empirical measure $d\lambda^N$ of the Lyapunov exponents converges in probability in the sense of moments to the measure

$$d\lambda^\infty := \frac{-e^{-z}}{S'_\rho(S_\rho^{-1}(e^{-z}))} \mathbf{1}_{[-\log S_\rho(-1), -\log S_\rho(0)]} dz.$$

The rescaled centered height function $M^{1/2}(\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)])$ converges in the sense of moments to the Gaussian random field on $[-\log S_\rho(-1), -\log S_\rho(0)]$ with covariance

$$K(t, s) = H\left(S_\rho^{-1}(e^{-t}) + 1, S_\rho^{-1}(e^{-s}) + 1\right) \frac{e^{-t}e^{-s}}{S'_\rho(S_\rho^{-1}(e^{-t}))S'_\rho(S_\rho^{-1}(e^{-s}))} + \delta(t - s)$$

for

$$H(u, w) := \frac{1}{M'_\rho(M_\rho^{-1}(u-1))M'_\rho(M_\rho^{-1}(w-1))(M_\rho^{-1}(u-1) - M_\rho^{-1}(w-1))^2} - \frac{1}{(u-w)^2}.$$

Proof. By Theorem 2.14, Lemma 2.7, and Theorem 4.5, $d\lambda^N$ converges in probability to the measure with moments

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(\lambda^N)] &= \oint \log(u/(u-1)) [-\log S_\rho(u-1)]^k \frac{du}{2\pi i} \\ &= \int_0^1 [-\log S_\rho(u-1)]^k du. \end{aligned}$$

Recall by [HL00, Theorem 4.4] that for a compactly supported measure $d\rho$ which is not a single atom, $S_\rho(z)$ is strictly decreasing on $[-1, 0]$. Therefore, we may change variables to obtain

$$\lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(\lambda^N)] = \int_{-\log S_\rho(-1)}^{-\log S_\rho(0)} z^k \frac{-e^{-z}}{S'_\rho(S_\rho^{-1}(e^{-z}))} dz,$$

where the values of $S_\rho(0)$ and $S_\rho(-1)$ are given by Lemma 3.11. These are the moments of $d\lambda^\infty$, which uniquely identify it since it is compactly supported.

For the central limit theorem, notice by integration by parts that

$$\begin{aligned}
& \int_{-\log S_{\bar{\rho}}(-1)}^{-\log S_{\bar{\rho}}(0)} \mathcal{H}_N(t) t^k dt \\
&= \frac{\mathcal{H}_N(-\log S_{\bar{\rho}}(0))[-\log S_{\bar{\rho}}(0)]^{k+1}}{k+1} - \frac{\mathcal{H}_N(-\log S_{\bar{\rho}}(-1))[-\log S_{\bar{\rho}}(-1)]^{k+1}}{k+1} - \int_{-\log S_{\bar{\rho}}(-1)}^{-\log S_{\bar{\rho}}(0)} \mathcal{H}'_N(t) \frac{t^{k+1}}{k+1} dt \\
&= \frac{\mathcal{H}_N(-\log S_{\bar{\rho}}(0))[-\log S_{\bar{\rho}}(0)]^{k+1}}{k+1} - \frac{\mathcal{H}_N(-\log S_{\bar{\rho}}(-1))[-\log S_{\bar{\rho}}(-1)]^{k+1}}{k+1} - \frac{1}{k+1} p_{k+1}(\lambda^N),
\end{aligned}$$

which means that

$$H_k := \int_{-\log S_{\bar{\rho}}(-1)}^{-\log S_{\bar{\rho}}(0)} M^{1/2}(\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) t^k dt = -\frac{M^{1/2}}{k+1} \left(p_{k+1}(\lambda^N) - \mathbb{E}[p_{k+1}(\lambda^N)] \right).$$

By Theorem 2.15, Lemma 2.7, and Theorem 4.5, we conclude that H_k are asymptotically Gaussian with covariance

$$\begin{aligned}
\lim_{N \rightarrow \infty} \text{Cov}(H_k, H_l) &= \oint \oint \log(u/(u-1)) \log(w/(w-1)) [-\log S_{\bar{\rho}}(u-1)]^k [-\log S_{\bar{\rho}}(w-1)]^l F^{(1,1)}(u, w) \frac{du}{2\pi i} \frac{dw}{2\pi i} \\
&\quad + \oint \log(u/(u-1)) [-\log S_{\bar{\rho}}(u-1)]^{k+l} \left[-\frac{S'_{\bar{\rho}}(u-1)}{S_{\bar{\rho}}(u-1)} \right] \frac{du}{2\pi i} \\
&= \int_0^1 \int_0^1 [-\log S_{\bar{\rho}}(u-1)]^k [-\log S_{\bar{\rho}}(w-1)]^l H(u, w) du dw - \int_0^1 [-\log S_{\bar{\rho}}(u-1)]^{k+l} \frac{S'_{\bar{\rho}}(u-1)}{S_{\bar{\rho}}(u-1)} du \\
&= \int_{-\log S_{\bar{\rho}}(-1)}^{-\log S_{\bar{\rho}}(0)} \int_{-\log S_{\bar{\rho}}(-1)}^{-\log S_{\bar{\rho}}(0)} t^k s^l K(t, s) dt ds
\end{aligned}$$

for

$$\begin{aligned}
H(u, w) &:= \frac{1}{M'_{\bar{\rho}}(M_{\bar{\rho}}^{-1}(u-1)) M'_{\bar{\rho}}(M_{\bar{\rho}}^{-1}(w-1)) (M_{\bar{\rho}}^{-1}(u-1) - M_{\bar{\rho}}^{-1}(w-1))^2} - \frac{1}{(u-w)^2} \\
K(t, s) &:= H\left(S_{\bar{\rho}}^{-1}(e^{-t}) + 1, S_{\bar{\rho}}^{-1}(e^{-s}) + 1\right) \frac{e^{-t} e^{-s}}{S'_{\bar{\rho}}(S_{\bar{\rho}}^{-1}(e^{-t})) S'_{\bar{\rho}}(S_{\bar{\rho}}^{-1}(e^{-s}))} + \delta(t-s),
\end{aligned}$$

which yields the desired claim. $\square$

Remark. The law of large numbers for Lyapunov exponents shown in Theorem 4.14 agrees with the previous results of [New86b, Theorem 2.11], [Kar08, Theorem 2], and [Tuc10, Theorem 5.1].

Remark. By Theorem 4.14, the fluctuations of Lyapunov exponents have a white noise component when $M, N \rightarrow \infty$, while by Corollary 4.9, they form a log-correlated Gaussian field when M stays finite. We demonstrate a formal limit transition between these cases. For M finite, suppose that all $d\rho_i$ are identical and equal to a non-atomic measure $d\mu$ in Corollary 4.9. Recall by [HL00, Theorem 4.4] that this means $S_{\bar{\mu}}(z)$ is strictly decreasing on $[-1, 0]$. Let $d\bar{\rho}^M = d\bar{\mu}^{\boxtimes M}$, and let $d\lambda^M$ be the empirical measure of the Lyapunov exponents, which is the pushforward of the corresponding $d\rho^M$ under the map $x \mapsto \frac{1}{M}x$. Applying [Tuc10, Theorem 5.1] and adjusting for a normalization factor of 2, we see that as $M \rightarrow \infty$, the measure $d\lambda^M$ converges to

$$d\lambda^\infty := \frac{-e^{-z}}{S'_{\bar{\mu}}(S_{\bar{\mu}}^{-1}(e^{-z}))} \mathbf{1}_{[-\log S_{\bar{\mu}}(-1), -\log S_{\bar{\mu}}(0)]} dz,$$

which coincides with the result of Theorem 4.14. Now define the height function

$$\mathcal{H}_{N,M}(t) := \#\{\rho_i^M \leq t\}.$$

By Corollary 4.9, its centered version $\mathcal{H}_{N,M}(t) - \mathbb{E}[\mathcal{H}_{N,M}(t)]$ converges as $N \rightarrow \infty$ to the Gaussian random field on $\mathbb{R}$ with covariance

$$K^M(t, s) = -\frac{M}{2\pi^2} \log \left| \frac{M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{-t}) + \mathbf{i}\pi p^M(t)) - M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{-s}) + \mathbf{i}\pi p^M(s))}{M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{-t}) + \mathbf{i}\pi p^M(t)) - M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{-s}) - \mathbf{i}\pi p^M(s))} \right| \\ + \frac{M-1}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}^M}(e^{-t}) - M_{\tilde{\rho}^M}(e^{-s}) + \mathbf{i}\pi(p^M(t) - p^M(s))}{M_{\tilde{\rho}^M}(e^{-t}) - M_{\tilde{\rho}^M}(e^{-s}) + \mathbf{i}\pi(p^M(t) + p^M(s))} \right|,$$

where $p^M(s)$ is the density of $d\tilde{\rho}^M(s)$, which satisfies

$$q(t) := \lim_{M \rightarrow \infty} Mp^M(tM) = \frac{-e^{-t}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-t}))} \mathbf{1}_{[-\log S_{\tilde{\mu}}(-1), -\log S_{\tilde{\mu}}(0)]}.$$

The height function of $d\lambda^M$ is given by $\mathcal{H}_{N,M}(tM)$ and hence its limiting covariance of

$$M^{1/2} \left(\mathcal{H}_{N,M}(tM) - \mathbb{E}[\mathcal{H}_{N,M}(tM)] \right)$$

is given by $\lim_{M \rightarrow \infty} MK^M(tM, sM)$. To analyze this limit, we consider the limits of the Stieltjes transform.

Lemma 4.15. For $r \in (-\log S_{\tilde{\mu}}(-1), -\log S_{\tilde{\mu}}(0))$, we have the following:

(a) uniformly on compact subsets we have

$$\lim_{\varepsilon \rightarrow 0^\pm} M_{\tilde{\rho}^M}(e^{(-r+\mathbf{i}\varepsilon)M}) = \left(S_{\tilde{\mu}}^{-1}(e^{-r}) + C_1 \right) \pm \mathbf{i} \left(\pi M^{-1} \frac{e^{-r}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-r}))} + C_2 \right),$$

where $C_1 = O(M^{-1})$ and $C_2 = O(M^{-2})$ are real;

(b) uniformly on compact subsets we have

$$\lim_{\varepsilon \rightarrow 0^\pm} \partial_r [M_{\tilde{\rho}^M}(e^{(-r+\mathbf{i}\varepsilon)M})] = \frac{e^{-r}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-r}))} + O(M^{-1}).$$

Proof. First, by definition we have

$$f(u) := \frac{1}{M} \log M_{\tilde{\rho}^M}^{-1}(u) = \log S_{\tilde{\mu}}(u) - \frac{1}{M} \log(1 + u^{-1}) \\ f'(u) = \frac{S_{\tilde{\mu}}'(u)}{S_{\tilde{\mu}}(u)} + \frac{1}{M} \frac{1}{u^2 + u}.$$

For (a), expanding in series we conclude that for $\varepsilon \neq 0$, we have

$$M_{\tilde{\rho}^M}(e^{(-r+\mathbf{i}\varepsilon)M}) = S_{\tilde{\mu}}^{-1}(e^{-r+\mathbf{i}\varepsilon}) + \frac{1}{M} \frac{e^{-r+\mathbf{i}\varepsilon}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-r+\mathbf{i}\varepsilon}))} \log \left(1 + \frac{1}{S_{\tilde{\mu}}^{-1}(e^{-r+\mathbf{i}\varepsilon})} \right) + O(M^{-2}).$$

Taking the limit as $\varepsilon \rightarrow 0^\pm$ yields the result. For (b), we find that for $u = M_{\tilde{\rho}^M}(e^{(-r+\mathbf{i}\varepsilon)M})$ that

$$\lim_{\varepsilon \rightarrow 0^\pm} \partial_r [M_{\tilde{\rho}^M}(e^{(-r+\mathbf{i}\varepsilon)M})] = \lim_{\varepsilon \rightarrow 0^\pm} \left[\frac{S'(u)}{S(u)} + \frac{1}{M} \frac{1}{u^2 + u} \right]^{-1} = \frac{e^{-r}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-r}))} + O(M^{-1}). \quad \square$$

For t, s in a compact subset of $(-\log S_{\tilde{\mu}}(-1), -\log S_{\tilde{\mu}}(0))$, we notice that

$$MK^M(tM, sM) = K_1^M(t, s) + K_2^M(t, s)$$

for

$$K_1^M(t, s) := \lim_{\varepsilon \rightarrow 0^+} -\frac{M}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}^M}(e^{(-t+\mathbf{i}\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s+\mathbf{i}\varepsilon)M})}{M_{\tilde{\rho}^M}(e^{(-t+\mathbf{i}\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s-\mathbf{i}\varepsilon)M})} \right| \\ K_2^M(t, s) := \lim_{\varepsilon \rightarrow 0^+} -\frac{M^2}{2\pi^2} \left[\log \left| \frac{M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+\mathbf{i}\varepsilon)M})) - M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-s+\mathbf{i}\varepsilon)M}))}{M_{\tilde{\rho}^M}(e^{(-t+\mathbf{i}\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s+\mathbf{i}\varepsilon)M})} \right| \right. \\ \left. - \log \left| \frac{M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+\mathbf{i}\varepsilon)M})) - M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-s-\mathbf{i}\varepsilon)M}))}{M_{\tilde{\rho}^M}(e^{(-t+\mathbf{i}\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s-\mathbf{i}\varepsilon)M})} \right| \right].$$

If $|t - s| > M^{-1/2}$, applying Lemma 4.15 we see that

$$\begin{aligned}
MK^M(tM, sM) &= \lim_{\varepsilon \rightarrow 0^+} \left(-\frac{M^2}{2\pi^2} \log \left| \frac{M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M})) - M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-s+i\varepsilon)M}))}{M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M})) - M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-s-i\varepsilon)M}))} \right| \right. \\
&\quad \left. + \frac{M^2 - M}{2\pi^2} \log \left| \frac{M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s+i\varepsilon)M})}{M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s-i\varepsilon)M})} \right| \right) \\
&= -\frac{M^2}{2\pi^2} \log \left| \frac{M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})) - M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})) + B_1 + \mathbf{i}\pi M^{-1}(\frac{q(t)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})))}) - \frac{q(s)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})))}) + B_2)}{M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})) - M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})) + B_1 + \mathbf{i}\pi M^{-1}(\frac{q(t)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})))}) + \frac{q(s)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})))}) + B_3)} \right| \\
&\quad + \frac{M^2 - M}{2\pi^2} \log \left| \frac{S_{\tilde{\mu}}^{-1}(e^{-t}) - S_{\tilde{\mu}}^{-1}(e^{-s}) + C_1 + \mathbf{i}\pi M^{-1}(q(t) - q(s) + C_2)}{S_{\tilde{\mu}}^{-1}(e^{-t}) - S_{\tilde{\mu}}^{-1}(e^{-s}) + C_1 + \mathbf{i}\pi M^{-1}(q(t) + q(s) + C_3)} \right| \\
&= -\frac{M^2}{4\pi^2} \log \left(1 - \frac{4\pi^2 M^{-2} \frac{q(t)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})))} \frac{q(s)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})))}}{(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})) - M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})))^2} + O(M^{-3}) \right) \\
&\quad + \frac{M^2 - M}{4\pi^2} \log \left(1 - \frac{4\pi^2 M^{-2} q(t)q(s)}{(S_{\tilde{\mu}}^{-1}(e^{-t}) - S_{\tilde{\mu}}^{-1}(e^{-s}))^2} + O(M^{-3}) \right) \\
&= \frac{q(t)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})))} \frac{q(s)}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})))} \frac{1}{(M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-t})) - M_{\tilde{\mu}}^{-1}(S_{\tilde{\mu}}^{-1}(e^{-s})))^2} \\
&\quad - \frac{q(t)q(s)}{(S_{\tilde{\mu}}^{-1}(e^{-t}) - S_{\tilde{\mu}}^{-1}(e^{-s}))^2} + O(M^{-1}) \\
&= K(t, s) + O(M^{-1}),
\end{aligned}$$

where the real constants $B_1, B_2, B_3, C_1, C_2, C_3$ are all of size $O(M^{-1})$, and $K(t, s)$ is the kernel of Theorem 4.14. On the other hand, if $|t - s| \leq M^{-1/2}$, by Lemma 4.15 we have

$$\begin{aligned}
K_1^M(t, s) &= \lim_{\varepsilon \rightarrow 0^+} \frac{M}{2\pi^2} \log \left| 1 + \frac{M_{\tilde{\rho}^M}(e^{(-s+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(s-i\varepsilon)M})}{M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s+i\varepsilon)M})} \right| \\
&= \lim_{\varepsilon \rightarrow 0^+} \frac{M}{2\pi^2} \log \left| 1 + \frac{M_{\tilde{\rho}^M}(e^{(-s+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(s-i\varepsilon)M})}{(t-s)\partial_r[M_{\tilde{\rho}^M}(e^{(-r+i\varepsilon)M})]} \right| \\
&= \frac{M}{2\pi^2} \log \left| 1 + \frac{2\mathbf{i}\pi M^{-1} \frac{e^{-s}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-s}))} + \mathbf{i}D_1}{(t-s) \frac{e^{-s}}{S_{\tilde{\mu}}'(S_{\tilde{\mu}}^{-1}(e^{-s}))} + O((t-s)M^{-1} + (t-s)^2)} \right| \\
&= \frac{M}{4\pi^2} \log \left(1 + \frac{4\pi^2}{(t-s)^2 M^2} + O(M^{-1} + M^{-2}(t-s)^{-1}) \right),
\end{aligned}$$

where in the second line r lies in the interval between s and t , and in the third line $D_1 = O(M^{-2})$ is real. Similarly, applying a second order Taylor expansion for $M_{\tilde{\mu}}^{-1}$, for

$$D_1 = \frac{1}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M})))} \quad \text{and} \quad D_2 = \frac{M_{\tilde{\mu}}''(M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M})))}{M_{\tilde{\mu}}'(M_{\tilde{\mu}}^{-1}(M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M})))^3}$$

and

$$\Delta_1 = M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s+i\varepsilon)M}) \quad \text{and} \quad \Delta_2 = M_{\tilde{\rho}^M}(e^{(-t+i\varepsilon)M}) - M_{\tilde{\rho}^M}(e^{(-s-i\varepsilon)M}),$$

we have by application of Lemma 4.15 that

$$\begin{aligned}
K_2^M(t, s) &= -\frac{M^2}{2\pi^2} \lim_{\varepsilon \rightarrow 0^+} \log \left| \frac{D_1 - \frac{1}{2}D_2\Delta_1 + O(|\Delta_1|^2)}{D_1 - \frac{1}{2}D_2\Delta_2 + O(|\Delta_2|^2)} \right| \\
&= \frac{M^2}{2\pi^2} \log \left| 1 + \frac{-i\pi M^{-1} \frac{e^t}{S'_\mu(S_\mu^{-1}(e^{-t}))} \frac{M''_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))}{M'_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))^3} + iO(M^{-2})}{\frac{1}{M'_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))} + O(M^{-1})} \right| \\
&= \frac{M^2}{4\pi^2} \log \left(1 + \pi^2 M^{-2} \frac{e^{-2t}}{S'_\mu(S_\mu^{-1}(e^{-t}))^2} \frac{M''_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))^2}{M'_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))^4} + O(M^{-3}) \right) \\
&= \frac{1}{4} \frac{e^{-2t}}{S'_\mu(S_\mu^{-1}(e^{-t}))^2} \frac{M''_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))^2}{M'_\mu(M_\mu^{-1}(S_\mu^{-1}(e^{-t})))^4} + O(M^{-1}).
\end{aligned}$$

We conclude that for any compactly supported continuous function f , we have

$$\begin{aligned}
&\lim_{M \rightarrow \infty} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} MK^M(tM, sM) f(t, s) dt ds \\
&= \lim_{M \rightarrow \infty} \left[\int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \mathbf{1}_{|t-s| > M^{-1/2}} K(t, s) f(t, s) dt ds \right. \\
&\quad \left. + \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \mathbf{1}_{|t-s| \leq M^{-1/2}} \left(K_1^M(t, s) + K_2^M(t, s) \right) f(t, s) dt ds \right] \\
&= \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \mathbf{1}_{|t-s| > M^{-1/2}} \left(K(t, s) - \delta(t-s) \right) f(t, s) dt ds \\
&\quad + \lim_{M \rightarrow \infty} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-M^{-1/2}}^{M^{-1/2}} \mathbf{1}_{s+x \in [-\log S_\mu(-1), -\log S_\mu(0)]} K_1^M(s+x, s) f(s+x, s) dx ds.
\end{aligned}$$

Using the integral identity

$$\int_{-\infty}^{\infty} \log \left(1 + \frac{4\pi^2}{y^2} \right) dy = 4\pi^2,$$

we notice that

$$\begin{aligned}
&\lim_{M \rightarrow \infty} \int_{-M^{-1/2}}^{M^{-1/2}} \mathbf{1}_{s+x \in [-\log S_\mu(-1), -\log S_\mu(0)]} K_1^M(s+x, s) f(s+x, s) dx \\
&= \frac{M}{4\pi^2} \int_{-M^{-1/2}}^{M^{-1/2}} \mathbf{1}_{s+x \in [-\log S_\mu(-1), -\log S_\mu(0)]} \log \left(1 + \frac{4\pi^2}{x^2 M^2} + O(M^{-1} + M^{-2}x^{-1}) \right) f(s+x, s) dx \\
&= \lim_{M \rightarrow \infty} \frac{1}{4\pi^2} \int_{-M^{-1/2}}^{M^{-1/2}} \mathbf{1}_{s+yM^{-1} \in [-\log S_\mu(-1), -\log S_\mu(0)]} \log \left(1 + \frac{4\pi^2}{y^2} + O(M^{-1} + M^{-1}y^{-1}) \right) f(s+yM^{-1}, s) dy \\
&= f(s, s) + o(1).
\end{aligned}$$

Substituting this into the previous expression implies that

$$\lim_{M \rightarrow \infty} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} MK^M(tM, sM) f(t, s) dt ds = \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} \int_{-\log S_\mu(-1)}^{-\log S_\mu(0)} K(t, s) f(t, s) dt ds.$$

This gives a limit transition between log-correlated Gaussian fields with covariance $K^M(t, s)$ and the Gaussian field with covariance $K(t, s)$, which has a white noise component.

4.5. Lyapunov exponents for Jacobi and Ginibre. We now consider Lyapunov exponents and their fluctuations for Jacobi and Ginibre matrices.

Theorem 4.16. If the X_N^i are drawn from the Jacobi ensemble with parameters α and R so that $\alpha/N = \hat{\alpha} + O(N^{-1})$ and $R/N = \hat{R} + O(N^{-1})$, then as $M, N \rightarrow \infty$, the empirical measure $d\lambda^N$ of the Lyapunov exponents converges in probability in the sense of moments to the measure

$$d\lambda^\infty(z) = \frac{e^z(\hat{\alpha} + \hat{R} - 1)}{(1 - e^z)^2} \mathbf{1}_{[\log \frac{\hat{\alpha}}{\hat{\alpha} + \hat{R}}, \log \frac{\hat{\alpha} + 1}{\hat{\alpha} + \hat{R} + 1}]} dz.$$

For the height function

$$\mathcal{H}_N(t) := \#\{\lambda_i^N \leq t\},$$

the rescaled recentered height function $M^{1/2}(\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)])$ converges in the sense of moments to the Gaussian white noise on $[\log \frac{\hat{\alpha}}{\hat{\alpha} + \hat{R}}, \log \frac{\hat{\alpha} + 1}{\hat{\alpha} + \hat{R} + 1}]$.

Proof. By Theorem 2.14, Lemma 2.7, and Theorem 4.6, $d\lambda^N$ converges in probability to the measure with moments

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(\lambda^N)] &= \oint \log(u/(u-1)) \left[\log \frac{\hat{\alpha} + u}{\hat{\alpha} + \hat{R} + u} \right]^k \frac{du}{2\pi i} \\ &= \int_0^1 \left[\log \frac{\hat{\alpha} + u}{\hat{\alpha} + \hat{R} + u} \right]^k du \\ &= \int_{\log \frac{\hat{\alpha}}{\hat{\alpha} + \hat{R}}}^{\log \frac{\hat{\alpha} + 1}{\hat{\alpha} + \hat{R} + 1}} z^k \frac{e^z(\hat{\alpha} + \hat{R} - 1)}{(1 - e^z)^2} dz, \end{aligned}$$

so the law of large numbers follows because $d\lambda^\infty$ is compactly supported.

For the central limit theorem, we see as in the proof of Theorem 4.14 that

$$H_k := M^{1/2} \int_{\log \frac{\hat{\alpha}}{\hat{\alpha} + \hat{R}}}^{\log \frac{\hat{\alpha} + 1}{\hat{\alpha} + \hat{R} + 1}} (\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)]) t^k dt$$

are asymptotically Gaussian with covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}(H_k, H_l) &= \oint \log(u/(u-1)) \left[\log \frac{\hat{\alpha} + u}{\hat{\alpha} + \hat{R} + u} \right]^{k+l} \frac{\hat{R}}{(\hat{\alpha} + u)(\hat{\alpha} + \hat{R} + u)} \frac{du}{2\pi i} \\ &= \int_0^1 \left[\log \frac{\hat{\alpha} + u}{\hat{\alpha} + \hat{R} + u} \right]^{k+l} \frac{\hat{R}}{(\hat{\alpha} + u)(\hat{\alpha} + \hat{R} + u)} du \\ &= \int_{\log \frac{\hat{\alpha}}{\hat{\alpha} + \hat{R}}}^{\log \frac{\hat{\alpha} + 1}{\hat{\alpha} + \hat{R} + 1}} t^{k+l} dt \end{aligned}$$

by Theorem 2.15, Lemma 2.7, and Theorem 4.6. $\square$

Theorem 4.17. If $X_N^i = (G_N^i)^* G_N^i$ with G_N^i from the Ginibre ensemble with parameter L so that $L/N = \gamma + O(N^{-1})$ for $\gamma > 1$, then as $M, N \rightarrow \infty$, the empirical measure $d\lambda^N$ of the Lyapunov exponents converges in probability in the sense of moments to the measure

$$d\lambda^\infty(z) = e^z \mathbf{1}_{[\log(\gamma-1), \log \gamma]} dz.$$

For the height function $\mathcal{H}_N(t) := \#\{\lambda_i^N \leq t\}$, the rescaled recentered height function $M^{1/2}(\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)])$ converges in the sense of moments to the Gaussian white noise on $[\log(\gamma-1), \log \gamma]$.

Proof. By Theorem 2.14, Lemma 2.7, and Theorem 4.7, $d\lambda^N$ converges in probability to the measure with moments

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(\lambda^N)] &= \oint \log(u/(u-1)) [\log(u + \gamma - 1)]^k \frac{du}{2\pi i} \\ &= \int_0^1 [\log(u + \gamma - 1)]^k du \\ &= \int_{\log(\gamma-1)}^{\log \gamma} z^k e^z dz, \end{aligned}$$

which implies the desired because $d\lambda^\infty$ is compactly supported.

For the central limit theorem, as in the proof of Theorem 4.14, we see that

$$H_k := M^{1/2} \int_{\log(\gamma-1)}^{\log \gamma} \left(\mathcal{H}_N(t) - \mathbb{E}[\mathcal{H}_N(t)] \right) t^k dt$$

are asymptotically Gaussian with covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}(H_k, H_l) &= \oint \log(u/(u-1)) [\log(u+\gamma-1)]^{k+l} \frac{1}{u+\gamma-1} \frac{du}{2\pi i} \\ &= \int_0^1 [\log(u+\gamma-1)]^{k+l} \frac{1}{u+\gamma-1} du \\ &= \int_{\log(\gamma-1)}^{\log \gamma} t^{k+l} dt \end{aligned}$$

by Theorem 2.15, Lemma 2.7, and Theorem 4.7. $\square$

Remark. While Theorem 4.17 does not apply to the case $\gamma = 1$, we extrapolate the results to this case to compare to the literature. We see that the law of large numbers in Theorem 4.17 agrees with the triangular law on the exponentials of the Lyapunov exponents shown in [IN92]. For the central limit theorem, in [ABK14] and [For15], it was shown for the M -fold product of $N \times N$ Ginibre matrices with finite N and large M , it was shown that the i^{th} Lyapunov exponent satisfies

$$\lambda_i^N = \left(\Psi(N-i+1) - \log N \right) + M^{-1/2} X_i + O(M^{-1}),$$

where $X_i \sim \mathcal{N}(0, \Psi'(N-i+1))$ and $\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function. We will take the formal limit as $M, N \rightarrow \infty$. For compactly supported smooth functions f_1, f_2 on $(-\infty, 0]$, we apply the expansion

$$f_j(\lambda_i^N) = f_j \left(\Psi(N-i+1) - \log N \right) + M^{-1/2} X_i f_j' \left(\Psi(N-i+1) - \log N \right) + O(M^{-1})$$

to find that

$$\begin{aligned} MCov \left(\sum_{i=1}^N f_1(\lambda_i^N), \sum_{j=1}^N f_2(\lambda_j^N) \right) &= \sum_{i,j=1}^N f_1' \left(\Psi(N-i+1) - \log N \right) f_2' \left(\Psi(N-j+1) - \log N \right) \text{Cov}(X_i, X_j) + O(M^{-1/2}) \\ &= \sum_{i=1}^N f_1' \left(\Psi(N-i+1) - \log N \right) f_2' \left(\Psi(N-i+1) - \log N \right) \Psi'(N-i+1) + O(M^{-1/2}). \end{aligned}$$

Recalling the expansion $\Psi(x) = \log(x) - \frac{1}{2x} + O(x^{-2})$, as $N \rightarrow \infty$ we have the expansions

$$\begin{aligned} \Psi(N-tN+1) - \log N &= \log(1-t) + O(N^{-1}) \\ \Psi'(N-tN+1) &= N^{-1} \frac{1}{1-t} + O(N^{-1}). \end{aligned}$$

This yields

$$\begin{aligned} MCov \left(\sum_{i=1}^N f_1(\lambda_i^N), \sum_{j=1}^N f_2(\lambda_j^N) \right) &= \int_0^1 f_1'(\log(1-t)) f_2'(\log(1-t)) \frac{1}{1-t} dt + O(N^{-1} + M^{-1}) \\ &= \int_{-\infty}^0 f_1'(x) f_2'(x) dx + O(N^{-1} + M^{-1}), \end{aligned}$$

which suggests formally that the height function at $\gamma = 1$ converges to white noise on $(-\infty, 0]$. This heuristic computation coincides with the extrapolation of Theorem 4.17 to $\gamma = 1$.

5. A 2-D GAUSSIAN FIELD FROM PRODUCTS OF RANDOM MATRICES

5.1. Statement of the result. Suppose that $Y_N^1, \dots, Y_N^M$ are $N \times N$ random matrices which are right unitarily invariant, and let $X_N^i := (Y_N^i)^*(Y_N^i)$. In the setting where $M \rightarrow \infty$ and $N \rightarrow \infty$, we study the joint distribution of the Lyapunov exponents of

$$X_{N,\alpha} := (Y_N^1 \cdots Y_N^{\lfloor \alpha M \rfloor})^* (Y_N^1 \cdots Y_N^{\lfloor \alpha M \rfloor}) \quad \text{for } \alpha \in (0, 1).$$

Define the Lyapunov exponents and empirical measure for $X_{N,\alpha}$ by

$$\lambda_i^{N,\alpha} := \frac{1}{M} \log \mu_i^\alpha \quad \text{and} \quad d\lambda^{N,\alpha} := \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i^{N,\alpha}},$$

where $\mu_1^\alpha \geq \dots \geq \mu_N^\alpha$ are the eigenvalues of $X_{N,\alpha}$. Define the height function by

$$\mathcal{H}_N(t, \alpha) := \#\{\lambda_i^{N,\alpha} \leq t\}.$$

In Theorem 5.1, whose proof is given in Section 5.3, we show that $\mathcal{H}_N(t, \alpha)$ has a limit shape with fluctuations forming a two-dimensional Gaussian field.

Theorem 5.1. If each X_N^i has deterministic spectrum satisfying Assumption 3.3 for a compactly supported non-atomic measure $d\rho$, then as $M, N \rightarrow \infty$, the log-spectral measure $d\lambda^{N,\alpha}$ converges in probability in the sense of moments to

$$d\lambda^{\infty,\alpha} := \frac{-\alpha^{-1}e^{-\alpha^{-1}z}}{S'_\rho(S_\rho^{-1}(e^{-\alpha^{-1}z}))} \mathbf{1}_{[-\alpha \log S_\rho(-1), -\alpha \log S_\rho(0)]} dz.$$

The rescaled centered height function $M^{1/2}(\mathcal{H}_N(t, \alpha) - \mathbb{E}[\mathcal{H}_N(t, \alpha)])$ converges in the sense of moments to the Gaussian random field on

$$D := \{(t, \alpha) \mid t \in [-\alpha \log S_\rho(-1), -\alpha \log S_\rho(0)]\}$$

with covariance

$$\begin{aligned} K(t, \alpha; s, \beta) &= \alpha^{-1} H\left(S_\rho^{-1}(e^{-\alpha^{-1}t}) + 1, S_\rho^{-1}(e^{-\beta^{-1}s}) + 1\right) \frac{e^{-\alpha^{-1}t} e^{-\beta^{-1}s}}{S'_\rho(S_\rho^{-1}(e^{-\alpha^{-1}t})) S'_\rho(S_\rho^{-1}(e^{-\beta^{-1}s}))} \\ &\quad + \alpha^{-1} \delta(\alpha^{-1}t - \beta^{-1}s) \\ H(u, w) &= \frac{1}{M'_\rho(M_\rho^{-1}(u-1)) M'_\rho(M_\rho^{-1}(w-1)) (M_\rho^{-1}(u-1) - M_\rho^{-1}(w-1))^2} - \frac{1}{(u-w)^2}. \end{aligned}$$

for $0 < \beta \leq \alpha < 1$.

Remark. We recall that the distribution $\delta(\alpha^{-1}t - \beta^{-1}s)$ is defined so that for any continuous function $f(t, s)$ on $\mathbb{R}^2$ we have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t, s) \delta(\alpha^{-1}t - \beta^{-1}s) dt ds = \alpha \beta \int_{-\infty}^{\infty} f(\alpha r, \beta r) dr.$$

Remark. Making the change of variables $\tilde{t} = \alpha^{-1}t$ and $\tilde{s} = \beta^{-1}s$, we find that the $\alpha^{-1}\delta(\alpha^{-1}t - \beta^{-1}s)$ piece of $K(t, \alpha; s, \beta)$ becomes $\beta\delta(\tilde{t} - \tilde{s})$, meaning that the slice of the Gaussian field in Theorem 5.1 along $D_t := \{(t\alpha, \alpha) \in D\}$ is a standard Brownian motion indexed by α .

5.2. Multilevel LLN and CLT via multivariate Bessel generating functions. Our proof of Theorem 5.1 is based on a generalization of Theorems 2.14 and 2.15 to measures whose multivariate Bessel generating functions are related. Let χ_N be N -tuples such that

$$\frac{1}{N} \sum_{i=1}^N \delta_{\chi_{N,i}/N} \rightarrow d\chi$$

for some compactly supported measure $d\chi$. Suppose that we have χ_N -smooth measures $d\tilde{\lambda}_N^1, \dots, d\tilde{\lambda}_N^M$, with corresponding multivariate Bessel generating functions $\psi_{\chi,N}^1(s), \dots, \psi_{\chi,N}^M(s)$ with respect to χ_N . Let $d\lambda_N^i$ be the pushforward of $d\tilde{\lambda}_N^i$ under the map $\lambda \mapsto \frac{1}{M}\lambda$. We assume the following condition on these measures.

Assumption 5.2. There exists a χ_N -smooth measure $d\sigma_N$ with corresponding multivariate Bessel generating function $\phi_{\chi,N}(s)$ so that

$$\psi_{\chi,N}^i(s) = \phi_{\chi,N}(s)^i \text{ for } i = 1, \dots, M.$$

Under this assumption, we have the following two-dimensional LLN and CLT.

Theorem 5.3. If Assumption 5.2 holds with $d\sigma_N$ being LLN-appropriate for χ_N , then for any $0 < \alpha < 1$, if x^α is distributed according to $d\lambda_N^{[\alpha M]}$, in probability we have

$$\lim_{N \rightarrow \infty} \frac{1}{N} p_k(x^\alpha) = \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(x^{[\alpha M]})] = \mathfrak{p}_k^\alpha := \alpha^k \oint \Xi(u) \Psi(u)^k \frac{du}{2\pi i},$$

where the u -contour encloses V_χ and lies within U . In addition, the random measures $\frac{1}{N} \sum_{i=1}^N \delta_{x_i^\alpha}$ converge in probability to a deterministic compactly supported measure $d\mu^\alpha$ with $\int x^k d\mu^\alpha(x) = \mathfrak{p}_k^\alpha$.

Proof. This follows from Theorem 2.14 after correcting for a difference in scaling. $\square$

Theorem 5.4. If Assumption 5.2 holds with $d\sigma_N$ being CLT-appropriate for χ_N , then if $\{x^\alpha \in \mathbb{R}^N \mid x_1^\alpha \geq \dots \geq x_N^\alpha\}$ is distributed according to $d\lambda_N^{[\alpha M]}$, the collection of random variables

$$\{M^{1/2}(p_k(x^\alpha) - \mathbb{E}[p_k(x^\alpha)])\}_{k \in \mathbb{N}, \alpha \in (0,1)}$$

converges in probability to a Gaussian variable with covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}\left(M^{1/2} p_k(x^\alpha), M^{1/2} p_l(x^\beta)\right) &= kl \alpha^{k-1} \beta^l \oint \oint \Xi(u) \Xi(w) \Psi(u)^{k-1} \Psi(w)^{l-1} \Lambda(u, w) \frac{du}{2\pi i} \frac{dw}{2\pi i} \\ &\quad + kl \alpha^{k-1} \beta^l \oint \Xi(u) \Psi(u)^{k+l-2} \Psi'(u) \frac{du}{2\pi i} \end{aligned}$$

for $1 > \alpha \geq \beta > 0$, where the u and w -contours enclose V_χ and lie within U , and the u -contour is contained inside the w -contour.

Proof. The proof is parallel to that of Theorem 2.15. We therefore describe the necessary modifications. Fix $1 > \alpha_l > \dots > \alpha_1 > 0$, $r_t \geq 1$, and $\{k_s^t\}_{1 \leq t \leq l, 1 \leq s \leq r_t}$. Let $\beta_i := \alpha_i - \alpha_{i-1}$, where we use the convention $\alpha_0 \equiv 0$. Notice that

$$\mathbb{E} \left[\prod_{t=1}^l \prod_{s=1}^{r_t} p_{k_s^t}(x^{\alpha_t}) \right]$$

is given by the evaluation at $s = \chi_N$ of the quantity

$$\begin{aligned} &\frac{\phi_{\chi,N}(s)^{M - [\alpha_l M]} D_{k_1^l} \dots D_{k_{r_l}^l} \phi_{\chi,N}(s)^{[\alpha_l M] - [\alpha_{l-1} M]} \dots D_{k_1^1} \dots D_{k_{r_1}^1} \phi_{\chi,N}(s)^{[\alpha_1 M]}}{\phi_{\chi,N}(s)^M} \\ &= \frac{D_{k_1^l} \dots D_{k_{r_l}^l} \phi_{\chi,N}(s)^{[\alpha_l M]}}{\phi_{\chi,N}(s)^{[\alpha_l M]}} \frac{D_{k_1^{l-1}} \dots D_{k_{r_{l-1}}^{l-1}} \phi_{\chi,N}(s)^{[\alpha_{l-1} M]}}{\phi_{\chi,N}(s)^{[\alpha_{l-1} M]}} \dots \frac{D_{k_1^1} \dots D_{k_{r_1}^1} \phi_{\chi,N}(s)^{[\alpha_1 M]}}{\phi_{\chi,N}(s)^{[\alpha_1 M]}}. \end{aligned}$$

As in Section 2.4, we may expand this quantity into the sum of terms associated with forests on the vertex set $\{(t, s)\}$. The coefficients of this linear combination are independent of N and are monomials in M and β_i whose exponent in M is at most $-W_L$, where W_L is the LLN weight of the term. Using (2.19) as in the proof of Theorem 2.15, the exponent of M in each term with non-zero coefficient to an n^{th} mixed cumulant is at most $1 - n$. For $n \geq 3$, this shows that any n^{th} cumulant is the sum of finitely many terms, which are of order at most $O(N^{1-n} M^{1-n/2}) = o(1)$, as needed.

It remains to compute the covariance; it is given by modifying (2.20) in the proof of Theorem 2.15 to

$$\begin{aligned} \text{Cov}\left(M^{1/2} p_k(x^\alpha), M^{1/2} p_l(x^\beta)\right) &= M^{-k-l+1} \oint \oint \left(\Xi(u) + M\alpha\Psi(u)\right)^k \\ &\quad \left(\Xi(w) + M\beta\Psi(w)\right)^l \left(\frac{1}{(u-w)^2} + M\beta\Lambda(w, u)\right) \frac{dw}{2\pi i} \frac{du}{2\pi i} + O(N^{-1}), \end{aligned}$$

where the u and w -contours enclose V_χ and lie within U , the u -contour is contained inside the w -contour, and we add a prefactor to $\Psi(u)$, $\Psi(w)$, and $\Lambda(w, u)$ to reflect which differential operator in

$$\frac{D_k}{\phi_{\chi, N}(s)^{\lfloor \alpha M \rfloor}} \phi_{\chi, N}(s)^{\lfloor \alpha M \rfloor} \frac{D_l}{\phi_{\chi, N}(s)^{\lfloor \beta M \rfloor}} \phi_{\chi, N}(s)^{\lfloor \beta M \rfloor}$$

each term originates from. Extracting the leading order term in M as in the proof of Theorem 2.15 yields the desired. $\square$

5.3. Proof of Theorem 5.1. First, for the law of large numbers, we see by Theorem 4.5, Lemma 2.7, and Theorem 5.3 that for $0 < \alpha < 1$, the rescaled log-spectral measure of $X_{N, \alpha}$ converges in probability in the sense of moments to the measure with moments

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \mathbb{E}[p_k(\lambda_N^{\lfloor \alpha M \rfloor})] &= \alpha^k \oint \log(u/(u-1)) [-\log S_{\tilde{\rho}}(u-1)]^k \frac{du}{2\pi i} \\ &= \int_0^1 [-\alpha \log S_{\tilde{\rho}}(u-1)]^k du \\ &= \int_{-\alpha \log S_{\tilde{\rho}}(-1)}^{-\alpha \log S_{\tilde{\rho}}(0)} z^k \frac{-\alpha^{-1} e^{-\alpha^{-1} z}}{S'_{\tilde{\rho}}(S_{\tilde{\rho}}^{-1}(e^{-\alpha^{-1} z}))} dz, \end{aligned}$$

where the final change of variables is legal because $d\rho$ is not a single atom, hence $S_{\tilde{\rho}}(z)$ is decreasing on $[-1, 0]$ by [HL00, Theorem 4.4]. Because this measure is compactly supported, convergence of moments implies convergence of random measures, and we see that

$$\lim_{N \rightarrow \infty} d\lambda^{N, \alpha} = \frac{-\alpha^{-1} e^{-\alpha^{-1} z}}{S'_{\tilde{\rho}}(S_{\tilde{\rho}}^{-1}(e^{-\alpha^{-1} z}))} \mathbf{1}_{[-\alpha \log S_{\tilde{\rho}}(-1), -\alpha \log S_{\tilde{\rho}}(0)]} dz.$$

For the central limit theorem, we see that

$$\begin{aligned} \int_{-\alpha \log S_{\tilde{\rho}}(-1)}^{-\alpha \log S_{\tilde{\rho}}(0)} \mathcal{H}_N(t, \alpha) t^k dt &= \frac{N[-\alpha \log S_{\tilde{\rho}}(0)]^{k+1}}{k+1} - \int_{-\alpha \log S_{\tilde{\rho}}(-1)}^{-\alpha \log S_{\tilde{\rho}}(0)} \mathcal{H}'_N(t, \alpha) \frac{t^{k+1}}{k+1} dt \\ &= \frac{N[-\alpha \log S_{\tilde{\rho}}(0)]^{k+1}}{k+1} - \frac{p_{k+1}(\lambda_N^{\lfloor \alpha M \rfloor})}{k+1}. \end{aligned}$$

By Theorem 4.5, Lemma 2.7, and Theorem 5.4, this implies that the quantities

$$H_{k, \alpha} := M^{1/2} \left(p_k(\lambda_N^{\lfloor \alpha M \rfloor}) - \mathbb{E}[p_k(\lambda_N^{\lfloor \alpha M \rfloor})] \right)$$

are jointly Gaussian with covariance

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}(H_{k, \alpha}, H_{l, \beta}) &= \alpha^k \beta^{l+1} \oint \oint \log(u/(u-1)) \log(w/(w-1)) [-\log S_{\tilde{\rho}}(u-1)]^k [-\log S_{\tilde{\rho}}(w-1)]^l F^{(1,1)}(u, w) \frac{du}{2\pi i} \frac{dw}{2\pi i} \\ &\quad + \alpha^k \beta^{l+1} \oint \log(u/(u-1)) [-\log S_{\tilde{\rho}}(u-1)]^{k+l} \left[-\frac{S'_{\tilde{\rho}}(u-1)}{S_{\tilde{\rho}}(u-1)} \right] \frac{du}{2\pi i} \\ &= \beta \int_0^1 \int_0^1 [-\alpha \log S_{\tilde{\rho}}(u-1)]^k [-\beta \log S_{\tilde{\rho}}(w-1)]^l F^{(1,1)}(u, w) du dw \\ &\quad - \beta \int_0^1 [-\alpha \log S_{\tilde{\rho}}(u-1)]^k [-\beta \log S_{\tilde{\rho}}(u-1)]^l \frac{S'_{\tilde{\rho}}(u-1)}{S_{\tilde{\rho}}(u-1)} du \\ &= \int_{-\alpha \log S_{\tilde{\rho}}(-1)}^{-\alpha \log S_{\tilde{\rho}}(0)} \int_{-\beta \log S_{\tilde{\rho}}(-1)}^{-\beta \log S_{\tilde{\rho}}(0)} t^k s^l K(t, \alpha; s, \beta) dt ds \end{aligned}$$

for $0 < \beta \leq \alpha < 1$ and

$$K(t, \alpha; s, \beta) = \alpha^{-1} H \left(S_{\tilde{\rho}}^{-1}(e^{-\alpha^{-1}t}) + 1, S_{\tilde{\rho}}^{-1}(e^{-\beta^{-1}s}) + 1 \right) \frac{e^{-\alpha^{-1}t} e^{-\beta^{-1}s}}{S'_{\tilde{\rho}}(S_{\tilde{\rho}}^{-1}(e^{-\alpha^{-1}t})) S'_{\tilde{\rho}}(S_{\tilde{\rho}}^{-1}(e^{-\beta^{-1}s}))} \\ + \alpha^{-1} \delta(\alpha^{-1}t - \beta^{-1}s) \\ H(u, w) = \frac{1}{M'_{\tilde{\rho}}(M_{\tilde{\rho}}^{-1}(u-1)) M'_{\tilde{\rho}}(M_{\tilde{\rho}}^{-1}(w-1)) (M_{\tilde{\rho}}^{-1}(u-1) - M_{\tilde{\rho}}^{-1}(w-1))^2} - \frac{1}{(u-w)^2}.$$

APPENDIX A. FLUCTUATIONS OF SUMS OF RANDOM MATRICES

In this appendix, we compute the distribution of fluctuations of sums of unitarily invariant random matrices. Let $A_N^i = U_i A U_i^*$ with U_i a $N \times N$ i.i.d. Haar unitary matrix and A diagonal with real spectrum μ^N , and define

$$X_{N,M} := A_N^1 + \cdots + A_N^M.$$

Proposition A.1. As $M \rightarrow \infty$, we have the convergence in distribution

$$\frac{1}{\sqrt{M}} \left(X_{N,M} - \mathbb{E}[X_{N,M}] \right) \xrightarrow{d} Y^N$$

where Y^N is a centered Hermitian Gaussian random matrix with covariance

$$\text{Cov}(Y_{ij}^N, Y_{i'j'}^N) := \mathbb{E}[Y_{ij}^N Y_{i'j'}^N] = \begin{cases} \frac{\Delta}{N+1} & i = j = i' = j' \\ -\frac{\Delta}{(N-1)(N+1)} & i = j, i' = j', i \neq i' \\ \frac{N}{(N-1)(N+1)} \Delta & i \neq j, i = j', j = i' \\ 0 & \text{otherwise} \end{cases},$$

where $\Delta = \frac{1}{N} \sum_{i=1}^N (\mu_i^N)^2 - \left(\frac{1}{N} \sum_{i=1}^N \mu_i^N \right)^2$.

Proof. By the ordinary central limit theorem, it suffices to check that the matrix elements of $M = U A U^*$ with U Haar unitary have the desired covariance. We have that

$$M_{ij} = \sum_{k=1}^N u_{ik} \mu_k^N \bar{u}_{jk},$$

which implies that

$$\text{Cov}(M_{ij}, M_{i'j'}) = \sum_{k,k'=1}^N \text{Cov}(u_{ik} \bar{u}_{jk}, u_{i'k'} \bar{u}_{j'k'}) \mu_k^N \mu_{k'}^N \\ \text{Cov}(M_{ij}, \bar{M}_{i'j'}) = \sum_{k,k'=1}^N \text{Cov}(u_{ik} \bar{u}_{jk}, u_{j'k'} \bar{u}_{i'k'}) \mu_k^N \mu_{k'}^N.$$

By [CM08, Lemma 14], we have that

$$\mathbb{E}[u_{ik} \bar{u}_{jk}] = \delta_{ij} \frac{1}{N} \\ \mathbb{E}[u_{ik} u_{i'k'} \bar{u}_{jk} \bar{u}_{j'k'}] = \frac{1}{(N-1)(N+1)} [\delta_{ij} \delta_{i'j'} + \delta_{ij'} \delta_{i'j} \delta_{kk'}] - \frac{1}{(N-1)N(N+1)} [\delta_{ij} \delta_{i'j'} \delta_{kk'} + \delta_{ij'} \delta_{i'j}].$$

Computing using these formulas yields the desired form of the covariance. $\square$

Remark. The matrix Y^N has the law of

$$\sqrt{\frac{N}{(N-1)(N+1)}} \Delta \cdot \left(X_N - \frac{1}{N} \text{Tr}(X_N) \cdot \text{Id}_N \right),$$

where X_N is distributed as GUE_N .

APPENDIX B. MULTIVARIATE BESSEL GENERATING FUNCTIONS AND THE CHOLESKY DECOMPOSITION

Let Y be a $N \times N$ random matrix which is invariant under right multiplication by U_N , and let $X = Y^*Y$ be the resulting Hermitian positive-definite random matrix. If $Y = UR$ with U unitary and R upper triangular is the QR decomposition of Y , then $X = R^*R$ is the Cholesky decomposition of X . The following result gives a geometric interpretation of the diagonal entries of R .

Lemma B.1. If $B_k \subset \mathbb{C}^N$ is a random complex k -dimensional unit ball and $B_{k-1} \subset B_k$ is a random $(k-1)$ -dimensional unit ball, then

$$R_{kk} \stackrel{d}{=} \frac{\text{vol}(Y(B_k))}{\text{vol}(Y(B_{k-1}))} \cdot \frac{\text{vol}(B_{k-1})}{\text{vol}(B_k)}.$$

Proof. Choose a nested sequence of random complex unit balls $B_1 \subset \dots \subset B_k \subset \mathbb{C}^N$ of increasing dimension. Notice that

$$\left(\frac{\text{vol}(Y(B_1))}{\text{vol}(B_1)}, \dots, \frac{\text{vol}(Y(B_k))}{\text{vol}(B_k)} \right) \stackrel{d}{=} (|Yu_1|, \dots, |Yu_1 \wedge \dots \wedge Yu_k|),$$

where $u_1, \dots, u_k$ are the first k columns of a Haar unitary matrix. Because Y is right unitarily invariant, we notice that $(Yu_1, \dots, Yu_k) \stackrel{d}{=} (y_1, \dots, y_k)$, where $y_1, \dots, y_k$ are the first k columns of Y . By the interpretation of QR-decomposition as Gram-Schmidt orthogonalization, we obtain that

$$(|Yu_1|, \dots, |Yu_1 \wedge \dots \wedge Yu_k|) \stackrel{d}{=} (R_{11}, \dots, R_{11} \cdots R_{kk}),$$

which implies the desired. $\square$

From Lemma B.1, we may see that the law of the diagonal entries of the Cholesky decomposition is multiplicative over products of random matrices.

Proposition B.2. Let $Y^1 = U^1 R^1$ and $Y^2 = U^2 R^2$ be the QR decompositions of independent right unitarily-invariant matrices, and let $Y^3 = Y^1 Y^2 = U^3 R^3$ be the QR decomposition of their product. We have the equality in law

$$R_{kk}^3 \stackrel{d}{=} R_{kk}^1 \cdot R_{kk}^2.$$

Proof. Let B_k be a random k -dimensional complex ball and $B_{k-1} \subset B_k$ a random $(k-1)$ -dimensional ball. Because Y^1 is right invariant, we have that

$$R_{kk}^3 \stackrel{d}{=} \frac{\text{vol}(Y^1 Y^2(B_k))}{\text{vol}(Y^1 Y^2(B_{k-1}))} \cdot \frac{\text{vol}(B_{k-1})}{\text{vol}(B_k)} = \left[\frac{\text{vol}(Y^2(B_k))}{\text{vol}(Y^2(B_{k-1}))} \cdot \frac{\text{vol}(B_{k-1})}{\text{vol}(B_k)} \right] \cdot \left[\frac{\text{vol}(Y^1(B'_k))}{\text{vol}(Y^1(B'_{k-1}))} \cdot \frac{\text{vol}(B_{k-1})}{\text{vol}(B_k)} \right],$$

where B'_k and B'_{k-1} are independent copies of B_k and B_{k-1} . Applying Lemma B.1 completes the proof. $\square$

We now give an interpretation of the multivariate Bessel generating function $\phi_X(s)$ in terms of this Cholesky decomposition which is similar to [KK16b, Lemma 5.3]. Together with Proposition B.2, this gives an independent geometric proof of Lemma 4.2.

Proposition B.3. If X is a Hermitian positive-definite unitarily-invariant random matrix with smooth log-spectral measure $d\mu$, then

$$\phi_X(s) = \mathbb{E} \left[\prod_{k=1}^N R_{kk}^{2(s_k - \rho_k)} \right],$$

where $X = R^*R$ is the Cholesky decomposition of X .

Proof. By the HCIZ integral of [HC57a, HC57b, IZ80], we have

$$\frac{\mathcal{B}(s, x)}{\mathcal{B}(\rho, x)} = \frac{\Delta(\rho) \det(e^{s_i x_j})}{\Delta(s) \Delta(e^x)} = \frac{\Delta(x)}{\Delta(e^x)} \int e^{\text{Tr}[SU \text{diag}(x) U^*]} d\text{Haar}_U,$$

where S is the diagonal matrix with diagonal entries $(s_1, \dots, s_N)$. Denote now by $\mathcal{O}_x$ the (coadjoint) orbit of the conjugation action of U_N on $\text{diag}(x_1, \dots, x_N)$ and by $d\omega_x$ the pushforward of $d\text{Haar}_U$ to $\mathcal{O}_x$ under the map $U \mapsto U \text{diag}(x) U^*$. Define the Gelfand-Tsetlin polytope GT_x by

$$\text{GT}_x := \{y_i^k, 1 \leq i \leq k, 1 \leq k \leq N \mid y^1 \prec \dots \prec y^N = x\},$$

where $y^k \prec y^{k+1}$ means that $y_{i+1}^{k+1} \leq y_i^k \leq y_i^{k+1}$ and define the map $\text{GT} : \mathcal{O}_x \rightarrow \text{GT}_x$ by

$$\text{GT}(Y) := (\text{eig}(Y_k))_{1 \leq k \leq N-1},$$

where Y_k denotes the principal $k \times k$ submatrix of Y . By [Bar01, Lemma 1.12 and Proposition 4.7] and [GN50], we have that

$$\text{GT}_*(d\omega_x) = \mathbf{1}_{y^1 \prec \dots \prec y^N = x} \frac{(N-1)! \dots 1!}{\Delta(x)} \prod_{k=1}^{N-1} \prod_{i=1}^k dy_i^k.$$

Applying this twice and applying the change of variables $z_i^k = e^{y_i^k}$ implies that

$$\begin{aligned} \frac{\mathcal{B}(s, x)}{\mathcal{B}(\rho, x)} &= \frac{(N-1)! \dots 1!}{\Delta(e^x)} \int_{y^1 \prec \dots \prec y^N = x} e^{\sum_{k=1}^N (s_k - s_{k+1}) \sum_{i=1}^k y_i^k} \prod_{k=1}^N \prod_{i=1}^k dy_i^k \\ &= \frac{(N-1)! \dots 1!}{\Delta(e^x)} e^{\sum_{i=1}^N x_i} \int_{z^1 \prec \dots \prec z^N = e^x} \prod_{k=1}^N \left(\prod_{i=1}^k z_i^k \right)^{s_k - s_{k+1} - 1} \prod_{k=1}^N \prod_{i=1}^k dz_i^k \\ &= e^{\sum_{i=1}^N x_i} \int \prod_{k=1}^N \det((Ue^x U^*)_k)^{s_k - s_{k+1} - 1} d\text{Haar}_U, \end{aligned}$$

where we adopt the convention that $s_{N+1} \equiv 0$. Now, for $Ue^x U^* = R^* R \in \mathcal{O}_{e^x}$, we see that $\det((Ue^x U^*)_k) = \det(R_k)^2 = \prod_{i=1}^k R_{ii}^2$, so rearranging the previous equation we conclude that

$$\frac{\mathcal{B}(s, x)}{\mathcal{B}(\rho, x)} = \int \prod_{k=1}^N R_{kk}^{2(s_k - \rho_k)} d\text{Haar}_U.$$

Therefore, if $d\mu_N$ is the log-spectral measure of X , we see that

$$\phi_X(s) = \int \frac{\mathcal{B}(s, x)}{\mathcal{B}(\rho, x)} d\mu(x) = \int \int \prod_{k=1}^N R_{kk}^{2(s_k - \rho_k)} d\text{Haar}_U d\mu(x) = \mathbb{E} \left[\prod_{k=1}^N R_{kk}^{2(s_k - \rho_k)} \right],$$

where $R^* R = X$ is the Cholesky decomposition of X . $\square$

We now combine Proposition B.3 with Theorem 4.5 to obtain a geometric interpretation of the S -transform. Let $d\lambda$ be a compactly supported measure on $(0, \infty)$, and for a fixed compact set $I \supset \text{supp } d\lambda$, let $\lambda^N \in \{x \in I^N \mid x_1 \geq \dots \geq x_N > 0\}$ be a sequence such that we have the weak convergence of measures

$$\frac{1}{N} \sum_{i=1}^N \delta_{\lambda^N} \rightarrow d\lambda.$$

Corollary B.4. Let X_N be the $N \times N$ unitarily invariant Hermitian random matrix with spectrum λ^N . For $t \in [0, 1]$, the log- S -transform of the measure $d\lambda$ is given by

$$-\log S_{d\lambda}(t-1) = \lim_{N \rightarrow \infty} \mathbb{E} [2 \log R_{[tN], [tN]}],$$

where $X_N = R^* R$ is the Cholesky decomposition of X_N .

Proof. First, by Proposition B.3, we find that

$$\partial_{s_k} \log \phi_{X_N}(\rho) = \partial_{s_k} \phi_{X_N}(\rho) = \int \partial_{s_k} \left[\prod_{k=1}^N R_{kk}^{2(s_k - \rho_k)} \right] \Big|_{s=\rho} d\text{Haar}_U = \int 2 \log R_{kk} d\text{Haar}_U,$$

where $X_N = R^* R$ is the Cholesky decomposition. Applying Theorem 4.5, we have uniformly in $k \in \{1, \dots, N\}$ that

$$\lim_{N \rightarrow \infty} \left| \int 2 \log R_{kk} d\text{Haar}_U + \log S_{d\lambda}(k/N - 1) \right| = 0,$$

from which the conclusion follows by continuity of the S -transform on $[-1, 0]$. $\square$

REFERENCES

- [ABK14] G. Akemann, Z. Burda, and M. Kieburg. Universal distribution of Lyapunov exponents for products of Ginibre matrices. *J. Phys. A*, 47(39):395202, 35, 2014.
- [ABK18] G. Akemann, Z. Burda, and M. Kieburg. From integrable to chaotic systems: Universal local statistics of Lyapunov exponents. *ArXiv e-prints*, 2018.
- [AKW13] G. Akemann, M. Kieburg, and L. Wei. Singular value correlation functions for products of Wishart random matrices. *J. Phys. A*, 46(27):275205, 22, 2013.
- [AM18] O. Arizmendi and J. A. Mingo. Second order cumulants: second order even elements and R -diagonal elements. *In preparation*, 2018.
- [Bar01] Y. Baryshnikov. GUEs and queues. *Probab. Theory Related Fields*, 119(2):256–274, 2001.
- [BB07] S. T. Belinschi and H. Bercovici. A new approach to subordination results in free probability. *J. Anal. Math.*, 101:357–365, 2007.
- [BBCF17] S. T. Belinschi, H. Bercovici, M. Capitaine, and M. Février. Outliers in the spectrum of large deformed unitarily invariant models. *Ann. Probab.*, 45(6A):3571–3625, 2017.
- [BD17] J. Breuer and M. Duits. Central limit theorems for biorthogonal ensembles and asymptotics of recurrence coefficients. *J. Amer. Math. Soc.*, 30(1):27–66, 2017.
- [BG56] Felix Aleksandrovich Berezin and Izrail Moiseevich Gel’fand. Some remarks on the theory of spherical functions on symmetric riemannian manifolds. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 5:311–351, 1956.
- [BG15a] A. Borodin and V. Gorin. General β -Jacobi corners process and the Gaussian free field. *Comm. Pure Appl. Math.*, 68(10):1774–1844, 2015.
- [BG15b] A. Bufetov and V. Gorin. Representations of classical Lie groups and quantized free convolution. *Geometric and Functional Analysis*, 25(3):763–814, 2015.
- [BG18a] A. Bufetov and V. Gorin. Fluctuations of particle systems determined by Schur generating functions. *Adv. Math.*, 338:702–781, 2018.
- [BG18b] A. Bufetov and V. Gorin. Fourier transform on high-dimensional unitary groups with applications to random tilings. 2018+. [arXiv:1712.09925](#).
- [BGS18] A. Borodin, V. Gorin, and E. Strahov. Product matrix processes as limits of random plane partitions. *Int. Math. Res. Not.*, to appear, 2018. [arXiv:1806.10855](#).
- [BH96] E. Brézin and S. Hikami. Correlations of nearby levels induced by a random potential. *Nuclear Phys. B*, 479(3):697–706, 1996.
- [BH97] E. Brézin and S. Hikami. Spectral form factor in a random matrix theory. *Phys. Rev. E* (3), 55(4):4067–4083, 1997.
- [Bia98] P. Biane. Processes with free increments. *Math. Z.*, 227(1):143–174, 1998.
- [BL17] C. Boutillier and Z. Li. Limit shape and height fluctuations of random perfect matchings on square-hexagon lattices. *ArXiv e-prints*, September 2017.
- [BLP16] Z. Bai, H. Li, and G. Pan. Central limit theorem for linear spectral statistics of large dimensional separable sample covariance matrices. *ArXiv e-prints*, November 2016.
- [Bor14a] A. Borodin. CLT for spectra of submatrices of Wigner random matrices. *Mosc. Math. J.*, 14(1):29–38, 170, 2014.
- [Bor14b] A. Borodin. CLT for spectra of submatrices of Wigner random matrices, II: Stochastic evolution. In *Random matrix theory, interacting particle systems, and integrable systems*, volume 65 of *Math. Sci. Res. Inst. Publ.*, pages 57–69. Cambridge Univ. Press, New York, 2014.
- [BV92] H. Bercovici and D. Voiculescu. Lévy-Hinčin type theorems for multiplicative and additive free convolution. *Pacific J. Math.*, 153(2):217–248, 1992.
- [Car14] F. Carlson. *Sur une class des séries de Taylor*. PhD thesis, Upsala, 1914.
- [CG18] C. Cuenca and V. Gorin. q -deformed character theory for infinite-dimensional symplectic and orthogonal groups. *ArXiv e-prints*, December 2018.
- [CM08] S. Chatterjee and E. Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008.
- [CMR17] F. Comets, G. R. Moreno Flores, and A. F. Ramirez. Random polymers on the complete graph. *ArXiv e-prints*, page arXiv:1707.01588, July 2017.
- [CMSS07] B. Collins, J. A. Mingo, P. Śniady, and R. Speicher. Second order freeness and fluctuations of random matrices. III. Higher order freeness and free cumulants. *Doc. Math.*, 12:1–70, 2007.
- [CO18] N. Coston and S. O’Rourke. Gaussian fluctuations for linear eigenvalue statistics of products of independent iid random matrices. *ArXiv e-prints*, September 2018.
- [Cop65] E. T. Copson. *Asymptotic expansions*. Cambridge Tracts in Mathematics and Mathematical Physics, No. 55. Cambridge University Press, New York, 1965.
- [CPS18] M. Chen, J. Pennington, and S. Schoenholz. Dynamical isometry and a mean field theory of RNNs: Gating enables signal propagation in recurrent neural networks. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 873–882, Stockholmssan, Stockholm Sweden, 10–15 Jul 2018. PMLR.
- [CPV93] A. Crisanti, G. Paladin, and A. Vulpiani. *Products of random matrices in statistical physics*, volume 104 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, 1993. With a foreword by Giorgio Parisi.
- [Cue18a] C. Cuenca. Asymptotic formulas for Macdonald polynomials and the boundary of the (q, t) -Gelfand-Tsetlin graph. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 14:Paper No. 001, 66, 2018.

- [Cue18b] C. Cuenca. Pieri integral formula and asymptotics of Jack unitary characters. *Selecta Math. (N.S.)*, 24(3):2737–2789, 2018.
- [Cue18c] C. Cuenca. Universal behavior of the corners of orbital beta processes. *ArXiv e-prints*, July 2018.
- [Dei17] P. Deift. Some open problems in random matrix theory and the theory of integrable systems. II. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 13:Paper No. 016, 23, 2017.
- [DJ18] M. Duits and K. Johansson. On mesoscopic equilibrium for linear statistics in Dyson’s Brownian motion. *Mem. Amer. Math. Soc.*, 255(1222):v+118, 2018.
- [DP12] I. Dumitriu and E. Paquette. Global fluctuations for linear statistics of β -Jacobi ensembles. *Random Matrices Theory Appl.*, 1(4):1250013, 60, 2012.
- [DP18] I. Dumitriu and E. Paquette. Spectra of overlapping Wishart matrices and the Gaussian free field. *Random Matrices Theory Appl.*, 7(2):1850003, 21, 2018.
- [Erd56] A. Erdélyi. *Asymptotic expansions*. Dover Publications, Inc., New York, 1956.
- [FK60] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.
- [For10] P. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010.
- [For13] P. J. Forrester. Lyapunov exponents for products of complex Gaussian random matrices. *J. Stat. Phys.*, 151(5):796–808, 2013.
- [For15] P. J. Forrester. Asymptotics of finite system Lyapunov exponents for some random matrix ensembles. *J. Phys. A*, 48(21):215205, 17, 2015.
- [GM17] V. Gorin and A. W. Marcus. Crystallization of random matrix orbits. *Int. Math. Res. Not.*, to appear, 2017.
- [GN50] I. M. Gelfand and M. A. Naïmark. *Unitarnye predstavleniya klassičeskikh grupp*. Trudy Mat. Inst. Steklov., vol. 36. Izdat. Nauk SSSR, Moscow-Leningrad, 1950.
- [GN15] A. Guionnet and J. Novak. Asymptotics of unitary multimatrix models: the Schwinger-Dyson lattice and topological recursion. *J. Funct. Anal.*, 268(10):2851–2905, 2015.
- [Gor14] V. Gorin. From alternating sign matrices to the Gaussian unitary ensemble. *Comm. Math. Phys.*, 332(1):437–447, 2014.
- [GP15] V. Gorin and G. Panova. Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. *Ann. Probab.*, 43(6):3052–3132, 2015.
- [GR84] H. Grauert and R. Remmert. *Coherent analytic sheaves*, volume 265 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1984.
- [HC57a] Harish-Chandra. Differential operators on a semisimple Lie algebra. *Amer. J. Math.*, 79:87–120, 1957.
- [HC57b] Harish-Chandra. Fourier transforms on a semisimple Lie algebra. *Amer. J. Math.*, 79:193–257, 1957.
- [Hel84] S. Helgason. *Groups and geometric analysis*, volume 113 of *Pure and Applied Mathematics*. Academic Press, Inc., Orlando, FL, 1984. Integral geometry, invariant differential operators, and spherical functions.
- [HL00] U. Haagerup and F. Larsen. Brown’s spectral distribution measure for R -diagonal elements in finite von Neumann algebras. *J. Funct. Anal.*, 176(2):331–367, 2000.
- [HL16] J. Huang and B. Landon. Local law and mesoscopic fluctuations of Dyson Brownian Motion for general β and potential. *ArXiv e-prints*, page arXiv:1612.06306, December 2016.
- [HN18] B. Hanin and M. Nica. Products of Many Large Random Matrices and Gradients in Deep Neural Networks. *arXiv e-prints*, page arXiv:1812.05994, December 2018.
- [Hua18] J. Huang. Law of large numbers and central limit theorems by Jack generating functions. *ArXiv e-prints*, July 2018.
- [IN92] M. Isopi and C. M. Newman. The triangle law for Lyapunov exponents of large random matrices. *Comm. Math. Phys.*, 143(3):591–598, 1992.
- [IZ80] C. Itzykson and J. Zuber. The planar approximation. ii. *J. Math. Phys.*, 21(3):411421, 1980.
- [Jia13] T. Jiang. Limit theorems for beta-Jacobi ensembles. *Bernoulli*, 19(3):1028–1046, 2013.
- [JL18] K. Johansson and G. Lambert. Gaussian and non-Gaussian fluctuations for mesoscopic linear statistics in determinantal processes. *Ann. Probab.*, 46(3):1201–1278, 2018.
- [Joh98] K. Johansson. On fluctuations of eigenvalues of random Hermitian matrices. *Duke Math. J.*, 91(1):151–204, 1998.
- [Joh04] K. Johansson. Determinantal processes with number variance saturation. *Comm. Math. Phys.*, 252(1-3):111–148, 2004.
- [Kar08] V. Kargin. Lyapunov exponents of free operators. *J. Funct. Anal.*, 255(8):1874–1888, 2008.
- [KK16a] M. Kieburg and H. Kösters. Exact relation between singular value and eigenvalue statistics. *Random Matrices Theory Appl.*, 5(4):1650015, 57, 2016.
- [KK16b] M. Kieburg and H. Kösters. Products of random matrices from polynomial ensembles. *ArXiv e-prints*, January 2016.
- [KZ14] A. B. J. Kuijlaars and L. Zhang. Singular values of products of Ginibre random matrices, multiple orthogonal polynomials and hard edge scaling limits. *Comm. Math. Phys.*, 332(2):759–781, 2014.
- [Li18a] Z. Li. Fluctuations of dimer heights on contracting square-hexagon lattices. *ArXiv e-prints*, September 2018.
- [Li18b] Z. Li. Schur function at generic points and limit shape of perfect matchings on contracting square hexagon lattices with piecewise boundary conditions. *ArXiv e-prints*, July 2018.
- [LQ18] Z. Ling and R. C. Qiu. Spectrum concentration in deep residual learning: a free probability approach. *arXiv e-prints*, page arXiv:1807.11694, July 2018.
- [LWW18] D.-Z. Liu, D. Wang, and Y. Wang. Lyapunov exponent, universality and phase transition for products of random matrices. *ArXiv e-prints*, September 2018.

- [Mac95] I. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.
- [MN18] S. Matsumoto and J. Novak. A moment method for invariant ensembles. *Electron. Res. Announc. Math. Sci.*, 2018.
- [MS06] J. A. Mingo and R. Speicher. Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces. *J. Funct. Anal.*, 235(1):226–270, 2006.
- [MS17] J. A. Mingo and R. Speicher. *Free probability and random matrices*, volume 35 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017.
- [MSS07] J. A. Mingo, P. Śniady, and R. Speicher. Second order freeness and fluctuations of random matrices. II. Unitary random matrices. *Adv. Math.*, 209(1):212–240, 2007.
- [New86a] C. M. Newman. The distribution of Lyapunov exponents: exact results for random matrices. *Comm. Math. Phys.*, 103(1):121–126, 1986.
- [New86b] C. M. Newman. Lyapunov exponents for some products of random matrices: exact expressions and asymptotic distributions. In *Random matrices and their applications (Brunswick, Maine, 1984)*, volume 50 of *Contemp. Math.*, pages 121–141. Amer. Math. Soc., Providence, RI, 1986.
- [NS97] A. Nica and R. Speicher. A “Fourier transform” for multiplicative functions on non-crossing partitions. *J. Algebraic Combin.*, 6(2):141–160, 1997.
- [Oko01] A. Okounkov. Infinite wedge and random partitions. *Selecta Math. (N.S.)*, 7(1):57–81, 2001.
- [Ose68] V. I. Oseledec. A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems. *Trudy Moskov. Mat. Obšč.*, 19:179–210, 1968.
- [OV96] G. Olshanski and A. Vershik. Ergodic unitarily invariant measures on the space of infinite Hermitian matrices. In *Contemporary mathematical physics*, volume 175 of *Amer. Math. Soc. Transl. Ser. 2*, pages 137–175. Amer. Math. Soc., Providence, RI, 1996.
- [Pil05] J. Pila. Note on Carlson’s theorem. *Rocky Mountain J. Math.*, 35(6):2107–2112, 2005.
- [PR04] D. Petz and J. Réffy. On asymptotics of large Haar distributed unitary matrices. *Period. Math. Hungar.*, 49(1):103–117, 2004.
- [PS11] L. Pastur and M. Shcherbina. *Eigenvalue distribution of large random matrices*, volume 171 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2011.
- [PSG17] J. Pennington, S. Schoenholz, and S. Ganguli. Resurrecting the sigmoid in deep learning through dynamical isometry: theory and practice. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 4785–4795. Curran Associates, Inc., 2017.
- [PSG18] J. Pennington, S. Schoenholz, and S. Ganguli. The emergence of spectral universality in deep networks. In *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain*, pages 1924–1932, 2018.
- [PV07] L. Pastur and V. Vasilchuk. On the law of addition of random matrices: covariance and the central limit theorem for traces of resolvent. In *Probability and mathematical physics*, volume 42 of *CRM Proc. Lecture Notes*, pages 399–416. Amer. Math. Soc., Providence, RI, 2007.
- [Rag79] M. S. Raghunathan. A proof of Oseledec’s multiplicative ergodic theorem. *Israel J. Math.*, 32(4):356–362, 1979.
- [Red16] N. K. Reddy. Lyapunov exponents and eigenvalues of products of random matrices. *ArXiv e-prints*, June 2016.
- [SMG14] A. M. Saxe, J. L. McClelland, and S. Ganguli. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. In *International Conference on Learning Representations (ICLR)*, 2014.
- [Tuc10] G. Tucci. Limits laws for geometric means of free random variables. *Indiana Univ. Math. J.*, 59(1):1–13, 2010.
- [TWJ⁺18] W. Tarnowski, P. Warchol, S. Jastrzebski, J. Tabor, and M. A. Nowak. Dynamical Isometry is Achieved in Residual Networks in a Universal Way for any Activation Function. *arXiv e-prints*, page arXiv:1809.08848, September 2018.
- [Vas16] V. Vasilchuk. On the fluctuations of eigenvalues of multiplicative deformed unitary invariant ensembles. *Random Matrices Theory Appl.*, 5(2):1650007, 28, 2016.
- [Voi85] D. Voiculescu. Symmetries of some reduced free product C^* -algebras. In *Operator algebras and their connections with topology and ergodic theory (Buşteni, 1983)*, volume 1132 of *Lecture Notes in Math.*, pages 556–588. Springer, Berlin, 1985.
- [Voi87] D. Voiculescu. Multiplication of certain noncommuting random variables. *J. Operator Theory*, 18(2):223–235, 1987.
- [Voi93] D. Voiculescu. The analogues of entropy and of Fisher’s information measure in free probability theory. I. *Comm. Math. Phys.*, 155(1):71–92, 1993.
- [XBSD⁺18] L. Xiao, Y. Bahri, J. Sohl-Dickstein, S. Schoenholz, and J. Pennington. Dynamical isometry and a mean field theory of CNNs: How to train 10,000-layer vanilla convolutional neural networks. In Jennifer Dy and Andreas Krause, editors, *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pages 5393–5402, Stockholmsmssan, Stockholm Sweden, 10–15 Jul 2018. PMLR.
- [ZJ99] P. Zinn-Justin. Adding and multiplying random matrices: a generalization of Voiculescu’s formulas. *Phys. Rev. E* (3), 59(5, part A):4884–4888, 1999.

V.G.: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WISCONSIN-MADISON, MADISON, WI, AND INSTITUTE FOR INFORMATION TRANSMISSION PROBLEMS OF RUSSIAN ACADEMY OF SCIENCES, MOSCOW, RUSSIA
E-mail address: `vadicgor@gmail.com`

Y.S.: DEPARTMENT OF STATISTICS, UNIVERSITY OF CHICAGO, CHICAGO, IL
E-mail address: `yisun@statistics.uchicago.edu`