Commun. Math. Phys. 394, 735–795 (2022)
Digital Object Identifier (DOI) https://doi.org/10.1007/s00220-022-04411-z

Communications in
Mathematical Physics

Matrix Addition and the Dunkl Transform at High Temperature

Florent Benaych-Georges[1,2], **Cesar Cuenca**[3], **Vadim Gorin**[4,5]

[1] Université Paris Descartes, Paris, France. E-mail: florent.benaych@gmail.com
[2] Capital Fund Management, Paris, France
[3] Harvard University, Cambridge, USA. E-mail: cesar.a.cuenk@gmail.com
[4] University of Wisconsin–Madison, Madison, USA. E-mail: vadicgor@gmail.com
[5] Institute for Information Transmission Problems of RAS, Moscow, Russia

Received: 16 August 2021 / Accepted: 22 April 2022
Published online: 2 June 2022 – © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract: We develop a framework for establishing the Law of Large Numbers for the eigenvalues in the random matrix ensembles as the size of the matrix goes to infinity simultaneously with the beta (inverse temperature) parameter going to zero. Our approach is based on the analysis of the (symmetric) Dunkl transform in this regime. As an application we obtain the LLN for the sums of random matrices as the inverse temperature goes to 0. This results in a one-parameter family of binary operations which interpolates between classical and free convolutions of the probability measures. We also introduce and study a family of deformed cumulants, which linearize this operation.

Contents

1. Introduction

1.1. Overview. This text comes out of two circles of ideas. On one side, we are interested in β-ensembles of random matrix theory, where $\beta = 1, 2, 4$ correspond to matrices with real/complex/quaternionic entries, but many distributions admit natural extensions to general real values of $\beta > 0$. The theoretical physics tradition refers to the β parameter as the inverse temperature. The matrices of interest are $N \times N$ and self-adjoint; we study the $N \to \infty$ asymptotic behavior of their eigenvalues on large scales. It was noticed by many authors (first, at classical $\beta = 1, 2, 4$ and later for all $\beta > 0$, see, e.g. [BAG97, J98, BoG15] for the results of the latter type) that in the global regime, when we deal with all eigenvalues together and describe the asymptotics of their empirical measures through Laws of Large Numbers and Central Limit Theorems, the only dependence of the answers on β is in simple normalization prefactors. In other words, the limits as $N \to \infty$ essentially do not depend on β, as long as $\beta > 0$ remains fixed. Recently, it was shown that the situation changes, if one varies β together with N in such a way that βN tends to a constant $2\gamma > 0$ as $N \to \infty$ (high-temperature regime). [ABG12, ABMV13, TT20] prove that for all classical ensembles of random matrices (Gaussian/Wigner, Laguerre/Wishart, and Jacobi/MANOVA) there is a different Law of Large Numbers in the high-temperature regime, and the resulting limit shapes non-trivially depend on the γ parameter. A subsequent wave produced many more results in the $\beta N \to 2\gamma$ asymptotic regime, such as the study of local statistics in [KS09, BP15, Pa18], or of central limit theorems in [NT18, HL21], or of the loop equations in [FM21], or of the spherical integrals in [MP21], or of the 2D systems in [AB19], or connections to Toda chain in [S19], or of dynamic versions in [NTT21]; this is very far from the complete list of results and we refer to the previously mentioned articles for further references.

From another side, a classical tool of the probability theory for establishing asymptotic theorems is by using the characteristic functions or Fourier transforms. In the last 10 years, a Fourier approach has been developed for the strongly correlated N-particle systems (with distributions of random-matrix type) in the series of papers [GP15, BuG15, BuG18, BuG19, MN18, Hu18, GSu18, Cu19, A20]. The central idea is to replace the exponents in the Fourier transform by symmetric functions of the representation-theoretic origin (such as Schur symmetric polynomials or multivariate Bessel functions) and to

further connect the partial derivatives of the logarithm of the new transform to the asymptotic behavior of the particle system (mostly, in the global regime) by using differential operators diagonalized by these symmetric functions.

In this article we develop a theory of integral transforms of N-tuples of real numbers (which should be thought of as eigenvalues of a random $N \times N$ matrix) using multivariate Bessel functions of general parameter $\theta = \frac{\beta}{2} > 0$ and generalizing conventional Fourier transform at $\theta = 0$; such transforms are also known as symmetric Dunkl transforms in the special functions literature, see [A17] for a review. We prove a very general theorem stating that the partial derivatives of the logarithms of our transforms at 0 have prescribed limits as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$ if and only if the associated random N-tuples satisfy a form of the Law of Large Numbers as $N \to \infty$, see Theorem 3.8. In our theory these partial derivatives play the same role as cumulants in classical probability and free cumulants in the free probability. We further develop a combinatorial theory of our new γ-cumulants in Theorems 3.10 and 3.11.

We present several applications of our theory:

- We recover previous results about Gaussian and Laguerre β-ensembles of random matrices as $\beta \to 0$, $N \to \infty$, $\beta N \to 2\gamma$, and recast them in the framework of γ-cumulants, see Sect. 4.1, Example 4.9, and Remark 4.11.
- We investigate eigenvalues of the sum of two independent self-adjoint matrices in the limit $\beta N \to 2\gamma$. We prove the Law of Large Numbers in this regime and encounter a new operation of γ-convolution, interpolating between usual convolution at $\gamma = 0$ and free (additive) convolution at $\gamma = \infty$, see Theorem 1.5.
- We obtain the Law of Large Numbers for ergodic Gibbs measures on the β-corners branching graph of [OV96,AN19] in the regime $\beta N \to 2\gamma$, see Theorem 4.10. The limits are infinitely-divisible with respect to γ-convolution.
- We find that each probability measure μ gives rise to a 1-parametric family of probability measures $\mu^{\tau,\gamma}$, $\tau \in [1, +\infty)$ which are $\beta N \to 2\gamma$ limits of empirical measures of spectra of $\lfloor N/\tau \rfloor \times \lfloor N/\tau \rfloor$ submatrix of $N \times N$ matrix whose spectrum approximates μ as $N \to \infty$. An intriguing property of the family is that all these measures are constructed from the same sequence of numbers, which are interpreted as the $\frac{\gamma}{\tau}$-cumulants of $\mu^{\tau,\gamma}$.

1.2. Addition of matrices as $\theta = \frac{\beta}{2} \to 0$. Rather than explaining our results in the most general and abstract setting, we focus on describing a particular application which was the original motivation for this work: the addition of random matrices. We start from a classical question. Let A and B be two self-adjoint $N \times N$ matrices with (real) eigenvalues $a_1 \le a_2 \le \cdots \le a_N$ and $b_1 \le b_2 \le \cdots \le b_N$, respectively. What can we say about the eigenvalues $c_1 \le c_2 \le \cdots \le c_N$ of the sum $C = A + B$?

The deterministic version of this problem asks to describe all possible values for $c_1 \le \cdots \le c_N$ if A and B are allowed to vary arbitrarily while preserving their eigenvalues. This question was first posed by Weyl [W12] in 1912 and it took the full XX century before it was completely resolved, see [KT] for a review. The answer is given by a convex set determined by the equality $\sum_{i=1}^N c_i = \sum_{i=1}^N a_i + \sum_{i=1}^N b_i$ (coming from $\mathrm{Trace}(C) = \mathrm{Trace}(A) + \mathrm{Trace}(B)$) and a large list of inequalities satisfied by $c_1, \ldots, c_N$: the simplest ones are well-known, for instance, $c_N \le a_N + b_N$, but there are many much more delicate relations.

The stochastic version of the same problem starts from random and independent matrices A and B. We assume that A is sampled from the uniform measure on the set of

all matrices with prescribed eigenvalues[1] $a_1 \le \cdots \le a_N$ and, similarly, B is a uniformly random matrix with eigenvalues $b_1 \le \cdots \le b_N$. Then the eigenvalues $c_1 \le \cdots \le c_N$ are random and we would like to obtain some description of them, with the most interesting questions pertaining to the situation of a very large N. The first asymptotic answer as $N \to \infty$ was obtained by Voiculescu in the context of the free probability theory.

Theorem 1.1 ([V91]; see also [Co03,CS06]). *Suppose that A and B are independent $N \times N$ uniformly random self-adjoint matrices with spectra $a_1(N) \le \cdots \le a_N(N)$ and $b_1(N) \le \cdots \le b_N(N)$, respectively, and let $c_1(N) \le \cdots \le c_N(N)$ be the (random) eigenvalues of $C = A + B$. Suppose that for two probability measures μ_A, μ_B, we have:*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{a_i(N)} = \mu_A, \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{b_i(N)} = \mu_B.$$

Then the random empirical measures $\frac{1}{N}\sum_{i=1}^{N} \delta_{c_i(N)}$ converge as $N \to \infty$ (weakly, in probability) to a deterministic measure $\mu_C := \mu_A \boxplus \mu_B$, which is called the free convolution of μ_A and μ_B.

In order to use this theorem, it is important to be able to efficiently describe the measure $\mu_A \boxplus \mu_B$. Let us briefly present two points of view on such description and refer to textbooks [NS06,MS17] for more details. The first point of view is analytic and it relies on the notion of the Voiculescu R-transform of a probability measure μ, defined through:

$$R_\mu(z) = (G_\mu(z))^{(-1)} - \frac{1}{z}, \qquad G_\mu(z) = \int_{\mathbb{R}} \frac{1}{z - x} \mu(dx),$$

where $G_\mu(z)$ is the Stieltjes transform of μ and $(G_\mu(z))^{(-1)}$ is the functional inverse. For a compactly supported μ, $R_\mu(z)$ is holomorphic in a complex neighborhood of 0. The measure $\mu_A \boxplus \mu_B$ is determined by:

$$R_{\mu_A \boxplus \mu_B}(z) = R_{\mu_A}(z) + R_{\mu_B}(z). \tag{1.1}$$

The relation (1.1) is a free probability version of the linearization of conventional convolution by logarithms of the characteristic functions: if ξ and η are independent random variables, then

$$\ln \mathbb{E} e^{\mathbf{i}t(\xi+\eta)} = \ln \mathbb{E} e^{\mathbf{i}t\xi} + \ln \mathbb{E} e^{\mathbf{i}t\eta}. \tag{1.2}$$

An alternative combinatorial approach to the free convolution uses free cumulants of a probability measure μ denoted κ_n^μ, $n = 1, 2, \ldots$. They are defined as certain explicit polynomials in the moments of the measure μ. Simultaneously, the free cumulants are coefficients of the Taylor-series expansion of $R_\mu(z)$ at the origin, so (1.1) gets restated as

$$\kappa_n^{\mu_A \boxplus \mu_B} = \kappa_n^{\mu_A} + \kappa_n^{\mu_B}, \quad n = 1, 2, \ldots. \tag{1.3}$$

This relation is a free probability version of the statement that conventional cumulants of a sum of independent random variables are sums of the cumulants of the summands.

Note that in Voiculescu's Theorem 1.1 we never specified, whether we deal with real symmetric, or complex Hermitian, or quaternionic Hermitian random matrices. And in

[1] Say, we deal with complex Hermitian matrices. Then this set is an orbit of the unitary group $U(N)$ under the action by conjugations, and the uniform measure on the orbit is the image of the Haar (uniform) measure on $U(N)$ with respect to this action.

fact, the theorem remains exactly the same in all these settings, which are usually referred as the $\beta = 1, 2, 4$ cases in the random matrix literature. What we would like to do is to go one step further and to extend the setting of the Theorem 1.1 to the general β setting. However, there is no (skew-)field of general real dimension $\beta > 0$, and therefore, there are no independent random matrices A and B over such field, which we could add. Hence, we first need to address a question:

Question. *What does it mean to add two independent self-adjoint β-random matrices A and B?*

Our answer to this question is based on the Fourier point of view on the addition of matrices. Suppose that $Q = [Q_{ij}]_{i,j=1}^N$ is a random real symmetric matrix. Its Fourier-Laplace transform is a function of another (deterministic) matrix X given by:

$$\chi_Q(X) = \mathbb{E}\exp\Big(\mathrm{Trace}(XQ)\Big) = \mathbb{E}\exp\Big(\sum_{i,j=1}^N x_{ij}Q_{ji}\Big). \tag{1.4}$$

Let us assume that the law of Q is invariant under conjugations by orthogonal matrices (which is the case for all three matrices A, B, and C in the Theorem 1.1). In addition assume that the matrix X is normal (i.e. $XX^* = X^*X$), which implies that X can be diagonalized by orthogonal conjugations[2]. In this situation, conjugating X and noting invariance of the trace, we see that $\chi_Q(X)$ is a function of the eigenvalues $x_1, \ldots, x_N$ of X and we can write it as $\chi_Q(x_1, \ldots, x_N)$.

If we specialize to the case when Q is a uniformly random real symmetric matrix with deterministic eigenvalues $q_1 \le \cdots \le q_N$, then χ_Q is known as a *multivariate Bessel function* at $\theta = \frac{\beta}{2} = \frac{1}{2}$:

$$\chi_Q(x_1, \ldots, x_N) = B_{(q_1,\ldots,q_N)}\big(x_1, \ldots, x_N;\ \tfrac{1}{2}\big). \tag{1.5}$$

Going further, the definition of χ_Q and linearity of the trace immediately imply that for independent conjugation-invariant matrices A and B we have

$$\chi_{A+B}(x_1, \ldots, x_N) = \chi_A(x_1, \ldots, x_N)\chi_B(x_1, \ldots, x_N). \tag{1.6}$$

Moreover, we can take (1.6) as a *definition* of $A + B$: the matrix $A + B$ is defined as a random $N \times N$ real symmetric matrix, whose law is invariant under orthogonal conjugations, and whose Fourier-Laplace transform is given by the right-hand side of (1.6).

The same argument can be given for complex Hermitian matrices and for quaternionic Hermitian matrices with the only difference being that the parameter of the Bessel functions in (1.5) changes to $\theta = 1$ and $\theta = 2$, respectively. But in fact, the multivariate Bessel functions $B_{(q_1,\ldots,q_N)}\big(x_1, \ldots, x_N;\ \theta\big)$ make sense for any real $\theta > 0$, see Sect. 2.2 for a formal definition. They are intimately connected to many topics, in particular, they are eigenfunctions of rational Calogero-Sutherland Hamiltonian and of (symmetric versions of) Dunkl operators; they are also limits of Jack and Macdonald symmetric polynomials. We are now ready to define the general β-analogue of addition of random matrices:

[2] If we know that Q is invariant under orthogonal conjugations and we know the values of $\chi_Q(X)$ for all normal X, then we can uniquely determine the law of Q. In fact it is sufficient to take X to be symmetric (or **i** times symmetric).

Definition 1.2. Fix $\theta = \frac{\beta}{2} > 0$. Given deterministic N-tuples of reals $\mathbf{a} = (a_1 \le \cdots \le a_N)$ and $\mathbf{b} = (b_1 \le \cdots \le b_N)$, we define a random N-tuple $\mathbf{c} = (c_1 \le \cdots \le c_N)$ by specifying its law through

$$\mathbb{E}\ B_{(c_1,\dots,c_N)}(x_1,\dots,x_N;\,\theta) = B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\,\theta)B_{(b_1,\dots,b_N)}(x_1,\dots,x_N;\,\theta), \qquad x_1,\dots,x_N \in \mathbb{C}. \tag{1.7}$$

We say that $\mathbf{c}$ is the eigenvalue distribution for the θ-sum of independent Hermitian matrices with spectra $\mathbf{a}$ and $\mathbf{b}$. We write $\mathbf{c} = \mathbf{a} +_\theta \mathbf{b}$.

For example, when $a_1 = \cdots = a_N = a$, the multivariate Bessel function is $B_{(a,\dots,a)}(x_1,\dots,x_N;\,\theta) = \exp(a(x_1 + \cdots + x_N))$. On the other hand, we have the identity

$$B_{(b_1+a,\dots,b_N+a)}(x_1,\dots,x_N;\,\theta) = \exp(a(x_1+\cdots+x_N))\,B_{(b_1,\dots,b_N)}(x_1,\dots,x_N;\,\theta), \tag{1.8}$$

as follows from Definition 2.5 below. So in the case that $\mathbf{a}_{\text{const}} = (a \le \cdots \le a)$ is the constant sequence, and $\mathbf{b} = (b_1 \le \cdots \le b_N)$ is arbitrary, then by comparing (1.7) and (1.8), we conclude that $\mathbf{c} = \mathbf{a}_{\text{const}} +_\theta \mathbf{b}$ is the Dirac delta mass at the point $(b_1 + a \le \cdots \le b_N + a)$. For more general sequences $\mathbf{a}$ we are not aware of similarly simple expressions for $\mathbf{a} +_\theta \mathbf{b}$.

Let us remark that the uniqueness of the law of $(c_1,\dots,c_N)$ defined through (1.7) is not hard to prove by expressing expectations of various test functions through expectations of multivariate Bessel functions.[3] In the existence part, there is a caveat. It is known that (1.7) defines $\mathbf{c}$ as a compactly supported generalized function (or distribution), see [T02], [A17, Section 3.6]. It is also straightforward to see that the total mass of distribution of $\mathbf{c}$ is 1 by inserting $x_1 = \cdots = x_N = 0$ into (1.7). However, the *positivity* of the law of $\mathbf{c}$, i.e. the fact there exists a *(probability) measure* on $\mathbf{c}$'s, such that (1.7) holds, is a well-known open question.

Conjecture 1.3. *Given any $\theta > 0$, and N-tuples $\mathbf{a} = (a_1 \le \cdots \le a_N)$, $\mathbf{b} = (b_1 \le \cdots \le b_N)$, there exists a probability measure on N-tuples $\mathbf{c} = (c_1 \le \cdots \le c_N)$ such that*

$$\mathbb{E}B_{(c_1,\dots,c_N)}(x_1,\dots,x_N;\,\theta) = B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\,\theta)B_{(b_1,\dots,b_N)}(x_1,\dots,x_N;\,\theta)$$

holds for any $x_1,\dots,x_N \in \mathbb{C}$. In other words, the distribution $\mathbf{c} = \mathbf{a} +_\theta \mathbf{b}$ is realized by a probability measure.

When $\theta = \frac{1}{2}, 1, 2$, the conjecture is known to be true, since we have a construction for $\mathbf{c}$ as eigenvalues of bona fide random matrices. In the limiting cases $\theta \to \infty$ and $\theta \to 0$ (here N is being fixed), the distribution $\mathbf{a} +_\theta \mathbf{b}$ turns into two explicit discrete probability measures, as we outline below. Conjecture 1.3, as well as its generalizations, have been mentioned in [S89, Conjecture 8.3], [R03a], [GM20, Conjecture 2.1], [M19, Section 1.2] and are believed to be true, yet we do not address it in our paper. Instead we state our results in such a way that they continue to hold even if the conjecture was wrong.

[3] For a reader who is not familiar with the theory of multivariate Bessel functions, we remark that at $N = 1$, $B_{(a)}(z;\,\theta) = \exp(az)$. Hence, choosing $z = \mathbf{i}t$, the Bessel functions turn into the exponents $\exp(\mathbf{i}at)$ and uniqueness turns into the well-known uniqueness of a measure with a given Fourier transform.

To sum up, the binary operation $+_\theta$ takes two deterministic N-tuples $\mathbf{a}$, $\mathbf{b}$ as input and outputs a distribution $\mathbf{c}$ on $\mathbb{R}^N$ (though conjecturally $\mathbf{c}$ is a random N-tuple). Even though there is no matrix interpretation for the operation $+_\theta$ for general values of $\theta > 0$, it is helpful to think of $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ as spectra of (nonexistent) self-adjoint $N \times N$ matrices. From our experience with random matrix theory, then the following question is natural: How does $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} +_\theta \mathbf{b}$ behave as $N \to \infty$? While this has not been written down in any published text, there are strong reasons to believe that as long as $\theta > 0$ is kept fixed, we get the same free convolution as in Theorem 1.1.[4] There are two boundary cases which need separate consideration: $\theta \to \infty$ and $\theta \to 0$. The former was addressed in [GM20], where it was proven that for fixed N, the $\theta \to \infty$ limit of $\mathbf{a} +_\theta \mathbf{b}$ is a deterministic operation known as finite free convolution; it was further shown in [Ma18] that as $N \to \infty$ we again recover the free convolution of Theorem 1.1. The final case $\theta \to 0$ turns out to be very different. The $\theta = 0$ version of multivariate Bessel function is a simple symmetric combination of exponents:

$$B_{(q_1,\ldots,q_N)}(x_1, \ldots, x_N;\, 0) = \frac{1}{N!} \sum_{\sigma \in S(N)} \prod_{i=1}^{N} \exp\bigl(x_i q_{\sigma(i)}\bigr), \tag{1.9}$$

where the sum goes over $N!$ different permutations of $\{1, 2, \ldots, N\}$. The formula (1.9) implies a transparent probabilistic interpretation: $\mathbf{c} = \mathbf{a} +_0 \mathbf{b}$ is obtained by choosing a permutation $\sigma \in S(N)$ uniformly at random and letting $(c_1, \ldots, c_N)$ be $(a_1 + b_{\sigma(1)}, \ldots, a_N + b_{\sigma(N)})$ rearranged in the increasing order. From this interpretation it is not hard to see that as $N \to \infty$ the operation $\mathbf{a} +_0 \mathbf{b}$ becomes the usual convolution of the empirical measures corresponding to $\mathbf{a}$ and to $\mathbf{b}$.

Hence, we see a discontinuity in the $N \to \infty$ behavior of the operation $(\mathbf{a}, \mathbf{b}) \mapsto (\mathbf{a} +_\theta \mathbf{b})$: at $\theta = 0$ the limit is described by the conventional convolution, while at $\theta > 0$ the limit is described by the free convolution. This motivates us to consider an intermediate scaling regime, in which θ goes to 0 as $N \to \infty$. This is the topic of the following Theorem 1.5, which is proven in Sect. 4.

Definition 1.4. We say that real random vectors $\mathbf{a}(N) = (a_1(N) \le a_2(N) \le \cdots \le a_N(N))$ converge as $N \to \infty$ in the sense of moments, if there exists a sequence of real numbers $\{m_k\}_{k \ge 1}$, such that for any $s = 1, 2, \ldots$ and any $k_1, k_2, \ldots, k_s \in \mathbb{Z}_{\ge 1}$, we have:

$$\lim_{N \to \infty} \mathbb{E}\left[\prod_{i=1}^{s} \left(\frac{1}{N} \sum_{j=1}^{N} \bigl(a_j(N)\bigr)^{k_i} \right) \right] = \prod_{i=1}^{s} m_{k_i}. \tag{1.10}$$

In this situation we write $\lim_{N \to \infty} \mathbf{a}(N) \overset{m}{=} \{m_k\}_{k \ge 1}$.

Note that (1.10) implies that the random empirical measures $\frac{1}{N} \sum_{i=1}^{N} \delta_{a_i(N)}$ converge as $N \to \infty$ weakly, in probability towards a deterministic measure with moments m_k, as long the moments problem associated with $\{m_k\}_{k \ge 1}$ has a unique solution. Also note that we can use Definition 1.4 in the situations where positivity of the distribution of $\mathbf{a}(N)$ is unknown: we may interpret $\mathbb{E}$ in (1.10) as the integral with respect to the distribution of $\mathbf{a}(N)$.

[4] The reasons are: widespread independence of the Law of Large Number from the value β for the random matrix β-ensembles, cf. [BAG97,J98,BoG15]; the same answer in Theorem 1.1 for three values $\theta = \frac{\beta}{2} = \frac{1}{2}, 1, 2$; existence of θ-independent observables for $\mathbf{a} +_\theta \mathbf{b}$, see [GM20, Theorem 1.1]; θ-independence in a discrete version of the same problem, see [Hu18].

Theorem 1.5. *Fix $\gamma > 0$ and suppose that $\theta > 0$ varies with N in such a way that $\theta N \to \gamma$ as $N \to \infty$. Take two sequences of independent random vectors $\mathbf{a}(N)$, $\mathbf{b}(N)$, $N = 1, 2, \dots$, such that*

$$\lim_{N\to\infty} \mathbf{a}(N) \overset{m}{=} \{m_k^{\mathbf{a}}\}_{k\geq 1}, \qquad \lim_{N\to\infty} \mathbf{b}(N) \overset{m}{=} \{m_k^{\mathbf{b}}\}_{k\geq 1}.$$

In addition, assume that $\mathbf{a}(N)$, $\mathbf{b}(N)$ satisfy the tail condition of Definition 2.8. Then

$$\lim_{\substack{N\to\infty \\ \theta N\to\gamma}} \left(\mathbf{a}(N) +_\theta \mathbf{b}(N)\right) \overset{m}{=} \{\tilde{m}_k\}_{k\geq 1},$$

where we call $\{\tilde{m}_k\}_{k\geq 1}$ the γ-convolution of $\{m_k^{\mathbf{a}}\}_{k\geq 1}$ and $\{m_k^{\mathbf{b}}\}_{k\geq 1}$ denoted through

$$\{\tilde{m}_k\}_{k\geq 1} = \{m_k^{\mathbf{a}}\}_{k\geq 1} \boxplus_\gamma \{m_k^{\mathbf{b}}\}_{k\geq 1}.$$

We further investigate the γ-convolution and establish the following properties:

1. There exist quantities called γ-*cumulants*, with the lth γ-cumulant $\kappa_l^{(\gamma)}$ being a homogeneous polynomial of degree l in the moments $m_1, \dots, m_l$ (where m_k is treated as a variable of degree k), such that for each $l = 1, 2, \dots$

$$\kappa_l^{(\gamma)}\left[\{m_k^{\mathbf{a}}\}_{k\geq 1} \boxplus_\gamma \{m_k^{\mathbf{b}}\}_{k\geq 1}\right] = \kappa_l^{(\gamma)}\left[\{m_k^{\mathbf{a}}\}_{k\geq 1}\right] + \kappa_l^{(\gamma)}\left[\{m_k^{\mathbf{b}}\}_{k\geq 1}\right]. \tag{1.11}$$

2. Each moment m_k can be expressed as a polynomial in $\kappa_l^{(\gamma)}$, $1 \leq l \leq k$, whose coefficients are explicit polynomials in γ with positive integer coefficients, see Theorem 3.10.
3. A generating function of the γ-cumulants $\kappa_l^{(\gamma)}$ is related to a generating function of the moments m_k through a simple relation, see Theorem 3.11.
4. As $\gamma \to 0$ the γ-convolution turns into the conventional convolution (i.e., if $\{m_k^{\mathbf{a}}\}_{k\geq 1}$ and $\{m_k^{\mathbf{b}}\}_{k\geq 1}$ are moments of two independent random variables, then $\lim_{\gamma\to 0} \{m_k^{\mathbf{a}}\}_{k\geq 1} \boxplus_\gamma \{m_k^{\mathbf{b}}\}_{k\geq 1}$ gives moments of their sum). After proper renormalization the γ-cumulants turn into conventional cumulants, see Sect. 8.1.
5. As $\gamma \to \infty$ the γ-convolution turns into the free convolution of Theorem 1.1. After proper renormalization the γ-cumulants turn into the free cumulants, see Sect. 8.2.

Remark 1.6. Given a probability measure μ with finite moments $\{m_k\}_{k\geq 1}$, we say that the corresponding $\{\kappa_l^{(\gamma)}\}_{l\geq 1}$ are the γ-*cumulants of* μ. It is known that the only probability measures with finitely many nonzero classical cumulants are Dirac delta masses and Gaussian distributions, see e.g. [L70, Thm. 7.3.3]. There is no such result in free probability (see [BV95, Thm. 2]): the semicircle distribution is a free probability analogue of the Gaussian distribution, but there are also very different measures with finitely many non-zero free cumulants. In our setting, the analogue of Gaussian/semicircle distributions are the measures for which only the first two γ-cumulants are nonzero, see Sect. 4.1 and Example 4.8. Therefore, a natural open question is whether there are more examples of probability measures with finitely many nonzero γ-cumulants.

Remark 1.7. We do not discuss in this text the *microscopic limits* of $\mathbf{a}(N) +_\theta \mathbf{b}(N)$ as $N \to \infty, \theta N \to \gamma$, i.e. the asymptotic questions in which individual eigenvalues remain visible in the limit. Yet, we expect to see the Poisson point process in the bulk of the spectrum, as hinted by general universality considerations and the $\theta \to 0$ asymptotic results in [KS09, AD14, BP15].

1.3. Law of large numbers through Bessel generating functions. Let us now outline the main technical tool, underlying the proof of Theorem 1.5 and other asymptotic results mentioned at the end of Sect. 1.1.

Suppose that $\mathbf{q} = (q_1 \le q_2 \le \cdots \le q_N)$ is a random N-tuple of reals. We define its Bessel generating function (BGF) through:

$$G_\theta(x_1, \dots, x_N; \mathbf{q}) = \mathbb{E}_{\mathbf{q}} \left[B_{(q_1,\dots,q_N)}(x_1, \dots, x_N;\ \theta) \right].$$

Our main result, Theorem 3.8, establishes an equivalence of the following two conditions for random sequences $\mathbf{q}(N) = (q_1(N), \dots, q_N(N))$ as $N \to \infty$ and $\theta \to 0$ in such a way that $\theta N \to \gamma$:

1. Partial derivatives of arbitrary order in x_1 of $\ln\bigl(G_\theta(x_1, \dots, x_N; \mathbf{q}(N))\bigr)$ at $(0, \dots, 0)$ converge to prescribed limits and partial derivatives in two (or more) different variables converge to 0.
2. Random vectors $\mathbf{q}(N)$ converge in the sense of moments, as in Definition 1.4.

The same theorem also establishes explicit polynomial formulas connecting the limiting value of the partial derivatives to the limiting values of the moments. The benefit of Theorem 3.8 is that it allows us to convert probabilistic information about $\mathbf{q}(N)$ into the analytic information about partial derivatives of its BGF and vice versa. For instance, Theorem 1.5 is then proven by three straightforward applications of Theorem 3.8: to $\mathbf{q}(N) = \mathbf{a}(N)$, to $\mathbf{q}(N) = \mathbf{b}(N)$, and to $\mathbf{q}(N) = \mathbf{a}(N) +_\theta \mathbf{b}(N)$. This and several other applications of Theorem 3.8 are detailed in Sect. 4.

Similar methods to the ones used in our proof of Theorem 3.8 have led to recent results in the literature, and even though our Theorem 3.8 bears resemblance to these other results, there are important differences. For example, in [BuG19] Bufetov and the third author developed a theory of Schur generating functions (SGF) for discrete N-particle systems as $N \to \infty$ (see also [Hu18] for an extension): they show that asymptotic information on partial derivatives of logarithms of SGF is in correspondence with asymptotic information on the moments in Law of Large Numbers as in Definition 1.4 and with covariances in a version of the Central Limit Theorem for global fluctuations. This is different from our Theorem 3.8: on the analytic side [BuG19] requires more refined control on partial derivatives and on the probabilistic side [BuG19] requires Central Limit Theorems in addition to Laws of Large Numbers.

In another similar framework related to multiplication of random matrices [GSu18] established a statement in one direction: control on partial derivatives implies the Law of Large Numbers and Central Limit Theorem, but in that framework a statement in the opposite direction remains out of reach.

Going further, we show in Section 9 that an analogue of Theorem 3.8 with fixed (rather than tending to 0) θ is wrong: there is no direct correspondence between partial derivatives of the logatithm of BGF and asymptotics of moments; one probably needs to use in such situation more complicated (and not yet understood) combinations of mixed partial derivatives in several variables. Thus, Theorem 3.8 is not an extension of the results of previous papers, but rather a brand new statement.

1.4. Connection to $\theta = \frac{\beta}{2} \to \infty$ limits. One intriguing aspect of general β random matrix theory is existence of dualities between parameters θ and $1/\theta$ (i.e. between β and $4/\beta$). In the theory of symmetric polynomials such a duality manifests through the existence of an automorphism of the algebra of symmetric functions, which transposes

the label of Jack symmetric polynomials and simultaneously inverts θ, see [S89, Section 3]. In the study of classical ensembles of random matrices the duality appears as a symmetry in expectations of power sums of eigenvalues, see, e.g., [DE06, Section 2.1], [FD16, Section 4.4], [F21], and references therein.

In our context, the duality suggests to look for a relation between $\theta \to 0$ limits of our paper and $\theta \to \infty$ limits. While this relation is not yet fully understood, we observe it in two forms.

First, the limit of the empirical measures of Gaussian β-ensembles as $\beta \to 0$, $N \to \infty$, $\beta N \to 2\gamma$ turns out to coincide with the orthogonality measure of the associated Hermite polynomials, see Remark 4.4. Simultaneously, the same polynomials play an important role in the study of centered fluctuations of Gaussian β-ensembles as $\beta \to \infty$ with N kept fixed, see [GK20, Section 4.5] and [AHV20].

Second, let us fix $N = d$ and send $\theta \to \infty$. [GM20, Theorem 1.2] claims that in this regime the operation $\mathbf{a} +_\theta \mathbf{b}$ turns into the finite free convolution, which is a deterministic binary operation on d-tuples of real numbers. Further, [AP18] introduced for each d a family of d finite free cumulants $\kappa^{\mathrm{ff}}_{1;d}, \kappa^{\mathrm{ff}}_{2;d}, \dots, \kappa^{\mathrm{ff}}_{d;d}$, which depend on a d-tuple of real numbers and play the same role for the finite free convolution, as our γ-cumulants play for the γ-convolution. Comparing the generating function of finite free cumulants from [AP18,Ma18], with the generating function of γ-cumulants of our Theorem 3.11, one sees[5] that upon setting $\gamma = -d$, they are very similar and only differ by normalizations, see Sect. 3.3 for more details. However, it is important to note that in our setting $\gamma > 0$, while in the setting of [AP18,Ma18], d is a positive integer and, thus, $-d$ is a negative integer. Hence, a correct point of view is that our γ-cumulants and the finite free cumulants are analytic continuations of each other. It would be interesting to see whether this observation can be used to produce new formulas for finite free cumulants along the lines of our Theorem 3.10.

2. Bessel Generating Functions

We define here the Bessel generating function of a probability measure on $\mathbb{R}^N$—this is a one-parameter generalization of the characteristic function (or Laplace transform) of a probability measure. The real parameter θ is assumed to be positive, with $\theta \to 0$ corresponding to the usual characteristic function. In this section, $\theta > 0$ remains fixed.

2.1. Difference and differential operators. We work with functions of N variables $x_1, \dots, x_N$. Denote the operator that permutes the variables x_i and x_j by $s_{i,j}$. For instance,

$$[s_{1,2} f](x_1, x_2, x_3, x_4, \dots, x_N) = f(x_2, x_1, x_3, x_4, \dots, x_N).$$

Define the *Dunkl operators* by

$$\mathcal{D}_i := \frac{\partial}{\partial x_i} + \theta \sum_{j: j \neq i} \frac{1}{x_i - x_j} \circ (1 - s_{i,j}), \quad i = 1, \dots, N. \tag{2.1}$$

[5] One should compare [AP18, (3.1), (4.2)] with our pair of equations (3.8) and notice that the conventions are slightly different: $(d)_n$ is a falling factorial in [AP18] and $(\gamma)_n$ is a rising factorial in our work. One can also directly compare the formulas for the first four cumulants of (3.3) and (3.6) with similar formulas above Corollary 4.3 in the journal version of [AP18]. We are grateful to Octavio Arizmendi and Daniel Perales for pointing this connection to us.

These operators were introduced in [D89]; see also [K97,R03b,EM10] for further studies. Their key property is commutativity:

$$\mathcal{D}_i\mathcal{D}_j = \mathcal{D}_j\mathcal{D}_i, \quad i, j = 1, \dots, N.$$

We often work with symmetrized versions of the Dunkl operators:

$$\mathcal{P}_k := (\mathcal{D}_1)^k + \cdots + (\mathcal{D}_N)^k, \quad k \in \mathbb{Z}_{\geq 1}.$$

Let $U \subseteq \mathbb{C}^N$ be any domain which is symmetric with respect to permutations of the axes. If f is a holomorphic function on U, then $\mathcal{D}_i f$ and $\mathcal{P}_k f$ are both well-defined and holomorphic on U.

We also need the *degree-lowering operators* $d_1, \dots, d_N$, which are defined on monomials by

$$d_i(x_1^{r_1} \cdots x_N^{r_N}) := \begin{cases} x_1^{r_1} \cdots x_i^{r_i-1} \cdots x_N^{r_N}, & \text{if } r_i \in \mathbb{Z}_{\geq 1}, \\ 0, & \text{if } r_i = 0, \end{cases} \tag{2.2}$$

and extended by linearity to the space of polynomials of N variables. They can be further extended to the ring of germs of analytic functions at the origin $(0, \dots, 0) \in \mathbb{C}^N$.

2.2. Multivariate Bessel functions. A central role in our studies is played by the simultaneous eigenfunctions of the operators $\mathcal{P}_k$ known as *multivariate Bessel functions*. They are given by very explicit formulas, which we describe next.

For each $N = 1, 2, \dots$, a *Gelfand–Tsetlin pattern of rank N* is an array $\{y_i^k\}_{1 \leq i \leq k \leq N}$ of real numbers satisfying $y_i^{k+1} \leq y_i^k \leq y_{i+1}^{k+1}$. Denote by $\mathcal{G}_N$ the space of all Gelfand–Tsetlin patterns of rank N.

Definition 2.1. Fix $\theta > 0$. The *θ-corners process with top row* $a_1 < \cdots < a_N$ is the probability distribution on the arrays $\{y_i^k\}_{1 \leq i \leq k \leq N} \in \mathcal{G}_N$, such that $y_i^N = a_i$, $i = 1, \dots, N$, and the remaining $N(N-1)/2$ coordinates have the density

$$\frac{1}{Z_{N;\theta}} \cdot \prod_{k=1}^{N-1} \left[\prod_{1 \leq i < j \leq k} (y_j^k - y_i^k)^{2-2\theta} \right] \cdot \left[\prod_{a=1}^{k} \prod_{b=1}^{k+1} |y_a^k - y_b^{k+1}|^{\theta-1} \right], \tag{2.3}$$

where $Z_{N;\theta}$ is the normalization constant:

$$Z_{N;\theta} = \left[\prod_{k=1}^{N} \frac{\Gamma(\theta)^k}{\Gamma(k\theta)} \right] \cdot \prod_{1 \leq i < j \leq N} (a_j - a_i)^{2\theta-1}. \tag{2.4}$$

Remark 2.2. By taking limits (in the space of probability measures on $\mathcal{G}_N$), we can allow equalities and extend the definition to arbitrary $a_1 \leq a_2 \leq \cdots \leq a_N$.

Remark 2.3. The distribution (2.3) is the joint law of eigenvalues of principal corners of Hermitian conjugation-invariant real/complex/quaternion matrices at $\theta = \frac{1}{2}, 1, 2$, respectively, see [Ne03]. This connects Definition 2.1 to the Laplace-Fourier point of view of Sect. 1.2.

Remark 2.4. The calculation of the normalization constant for a general $\theta > 0$ is contained in [Ne03] (the author there does far more general calculations; see also [Cu19, Lem. 2.1] for a short derivation of $Z_{N;\theta}$ from Anderson's integral identity [A91]).

Definition 2.5. The *multivariate Bessel function* $B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\theta)$ is defined as the following (partial) Laplace transform of the θ-corners process with top row $(a_1,\dots,a_N)$ from Definition 2.1:

$$B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\,\theta)=\mathbb{E}_{\{y_i^k\}}\left[\exp\left(\sum_{k=1}^{N}x_k\cdot\left(\sum_{i=1}^{k}y_i^k-\sum_{j=1}^{k-1}y_j^{k-1}\right)\right)\right]. \quad (2.5)$$

The function $B_{(a_1,\dots,a_N)}(x_1,\dots,x_N)$ is defined for any reals $a_1<\dots<a_N$ and any complex numbers $x_1,\dots,x_N$.

Often, we will abbreviate multivariate Bessel function as MBF.

It follows from the definition that

$$B_{(a_1,\dots,a_N)}(0,\dots,0;\theta)=1.$$

Our definition is called the *combinatorial formula* for the multivariate Bessel functions; to our knowledge, the formula (2.5) first appeared in [GuK02]. There are several alternative definitions of these functions. For example, from the algebraic combinatorics point of view, they can be defined as limits of (properly normalized) Jack symmetric polynomials. Then (2.5) is a limit of the combinatorial formulas for the Jack polynomials, cf. [OO97, Section 4].

The MBF $B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\theta)$, which was defined for ordered tuples $a_1<\dots<a_N$, can be extended to weakly ordered tuples $a_1\le\dots\le a_N$ by continuity: there is no singularity on the diagonals $a_i=a_j$. In fact, much more is true: $B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\theta)$ admits an analytic continuation on the $2N+1$ variables $a_1,\dots,a_N,x_1,\dots,x_N,\theta$, to an open subset of $\mathbb{C}^{2N+1}$ containing $\{(a_1,\dots,a_N,x_1,\dots,x_N,\theta)\in\mathbb{C}^{2N+1}\mid\Re\theta\ge 0\}$; see [O93]. In particular, for a fixed $\theta>0$, the MBF $B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\theta)$ is an entire function on the variables $a_1,\dots,a_N,x_1,\dots,x_N$.

Another important property is that the MBF $B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\,\theta)$ is *symmetric* in its arguments $x_1,\dots,x_N$—this is an immediate consequence of the fact that the MBFs are limits of properly normalized Jack symmetric polynomials, see e.g. (4.12) below. In the particular case $\theta=1$, the symmetry is also transparent from the following determinantal formula, which arises as the evaluation of the Harish-Chandra-Itzykson-Zuber (HCIZ) integral:

$$B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\,1)=1!\cdot 2!\cdots(N-1)!\cdot\frac{\det\left[e^{a_ix_j}\right]_{i,j=1}^{N}}{\prod_{i<j}(x_i-x_j)(a_i-a_j)}. \quad (2.6)$$

A link of MBF to the operators of Sect. 2.1 is given by the following statement.

Theorem 2.6 ([O93]). *For each $k=1,2,\dots$, and each N-tuple of reals $a_1\le a_2\le\dots\le a_N$,*

$$\mathcal{P}_k B_{(a_1,\dots,a_N)}=\left(\sum_{i=1}^{N}a_i^k\right)\cdot B_{(a_1,\dots,a_N)}. \quad (2.7)$$

2.3. Bessel generating functions. Let $\mathcal{M}_N$ be the convex set of Borel probability measures on ordered N-tuples $a_1 \le a_2 \le \cdots \le a_N$ of real numbers.

Definition 2.7. The *Bessel generating function* (or BGF) of $\mu \in \mathcal{M}_N$ is defined as a function of the variables $x_1, \dots, x_N$ given by:

$$G_\theta(x_1, \dots, x_N; \mu) := \int_{a_1 \le a_2 \le \cdots \le a_N} B_{(a_1,\dots,a_N)}(x_1, \dots, x_N; \theta)\mu(\mathrm{d}a_1, \dots, \mathrm{d}a_N). \tag{2.8}$$

Because the MBFs $B_{(a_1,\dots,a_N)}(x_1, \dots, x_N; \theta)$ are symmetric functions on the variables $x_1, \dots, x_N$, so is $G_\theta(x_1, \dots, x_N; \mu)$. Moreover,

$$G_\theta(0, \dots, 0; \mu) = 1,$$

as follows from μ being a probability measure and $B_{(a_1,\dots,a_N)}(0, \dots, 0; \theta) = 0$.

It will be important for us to assume that a BGF is defined in a *complex* neighborhood of $(0, \dots, 0)$. Unfortunately, this property fails for general measures, hence we need to restrict the class of measures that we deal with.[6]

Definition 2.8. We say that a measure $\mu \in \mathcal{M}_N$ is *exponentially decaying* with exponent $R > 0$, if

$$\int_{a_1 \le a_2 \le \cdots \le a_N} e^{NR \max_i |a_i|}\mu(\mathrm{d}a_1, \dots, \mathrm{d}a_N) < \infty. \tag{2.9}$$

Lemma 2.9. *If $\mu \in \mathcal{M}_N$ is exponentially decaying with exponent $R > 0$, then the integral (2.8) converges for all $(x_1, \dots, x_N)$ in the domain*

$$\Omega_R := \left\{(x_1, \dots, x_N) \in \mathbb{C}^N : |\Re x_i| < R,\ i = 1, \dots, N\right\},$$

and defines a holomorphic function in this domain.

Proof. Note that if $\{y_i^k\} \in \mathcal{G}_N$ satisfies $y_i^N = a_i$, $i = 1, \dots, N$, then due to interlacing inequalities, for each $k = 1, 2, \dots, N$ we have

$$\left|\sum_{i=1}^{k} y_i^k - \sum_{i=1}^{k-1} y_i^{k-1}\right| \le \max\left(|y_1^k|, |y_k^k|\right) \le \max_i |a_i|.$$

Hence, the integrand in the definition of the multivariate Bessel function (2.5) is upper bounded by

$$\exp\left(\sum_{j=1}^{N} |\Re x_j| \max_i |a_i|\right),$$

which implies

$$\left|B_{(a_1,\dots,a_N)}(x_1, \dots, x_N; \theta)\right| \le \exp\left(NR \max_i |a_i|\right), \quad (x_1, \dots, x_N) \in \Omega_R.$$

Hence, (2.9) implies convergence of the integral (2.8) in Ω_R.

[6] It is plausible that many of the results of our text extend to the situations where this restrictive condition fails.

It remains to check holomorphicity of (2.8) as a function of $x_1, \dots, x_N$. This readily follows from holomorphicity of $B_{(a_1,\dots,a_N)}(x_1, \dots, x_N; \theta)$. Indeed, $G_\theta(x_1, \dots, x_N; \mu)$ is continuous as a uniformly convergent integral of continuous functions. Thus, by Morera's theorem, the holomorphicity follows from vanishing of the integrals over closed contours. The latter vanishing can be deduced by swapping the integrations using the Fubini's theorem and using vanishing of the similar integrals for $B_{(a_1,\dots,a_N)}(x_1, \dots, x_N; \theta)$. □

The BGFs have recently been used in connection to problems in random matrix theory, see [Cu19,GSu18]. However, the BGF is not a new invention. The formula (2.8) is essentially the definition of (a symmetric version of) the *Dunkl transform*, a one-parameter generalization of the Fourier transform; this is a rich and well-studied subject, see e.g. the survey [A17] and references therein.

The next two propositions will be important in our developments.

Proposition 2.10. *Let $k \in \mathbb{Z}_{\geq 1}$ and let $\mu \in \mathcal{M}_N$ be an exponentially decaying measure. Then*

$$\left[\mathcal{P}_k\, G_\theta(x_1, \dots, x_N; \mu)\right]_{x_1=\dots=x_N=0} = \mathbb{E}_\mu \left[\sum_{j=1}^{N} (a_j)^k\right],$$

where $(a_1, \dots, a_N) \in \mathbb{R}^N$ is random and μ-distributed on the right-hand side.

Proof. We apply $\mathcal{P}_k$ to (2.8) under the sign of the integral, use the eigenrelation of Theorem 2.6 and the normalization $B_{(a_1,\dots,a_N)}(0, \dots, 0; \theta) = 1$. We can exchange the order between the operator $\mathcal{P}_k$ and the integral because μ is exponentially decaying. □

The following generalization of Proposition 2.10 is proved in the same way.

Proposition 2.11. *Let $k_1, \dots, k_s \in \mathbb{Z}_{\geq 1}$ and let $\mu \in \mathcal{M}_N$ be an exponentially decaying measure. Then*

$$\left(\prod_{i=1}^{s} \mathcal{P}_{k_i}\right) G_\theta(x_1, \dots, x_N; \mu)\Big|_{x_1=\dots=x_N=0} = \mathbb{E}_\mu \left[\prod_{i=1}^{s} \left(\sum_{j=1}^{N} (a_j)^{k_i}\right)\right]. \tag{2.10}$$

Observe that the pairwise commutativity of the Dunkl operators implies the pairwise commutativity of the operators $\mathcal{P}_k$, $k \in \mathbb{Z}_{\geq 1}$. As a result, the order of application of the operators $\mathcal{P}_{k_i}$ in the left-hand side of (2.10) does not matter.

2.4. Extension to distributions. Ultimately, we treat Bessel Generating Functions as a tool for studying symmetric probability measures on $\mathbb{R}^N$ (which can be identified with probability measures on ordered N-tuples $a_1 \leq a_2 \leq \dots \leq a_N$). One of the applications that we have in mind is to use them for the study of addition of independent general β random matrices. While it is conjectured that the spectrum of such sum should be described by a probability measure, it is not proven yet: we only rigorously know that the spectrum can be described as a generalized function or distribution (the technical problem is in proving positivity; see [T02], [A17, Section 3.6]). In order to avoid the necessity to rely on the positivity conjectures, we explain in this section that the framework of Bessel generating functions can be extended to objects more general than probability measures.

Let μ be a *distribution* on $\mathbb{R}^N$ with coordinates $(a_1, \ldots, a_N)$, i.e. μ is an element of the dual space to the space of compactly supported infinitely-differentiable *test-functions*.[7] μ is said to be symmetric if for any test-function f and any permutation σ:

$$\langle \mu, f(a_1, \ldots, a_N)\rangle = \langle \mu, f(a_{\sigma(1)}, \ldots, a_{\sigma(N)})\rangle,$$

where we use the notation $\langle \mu, f\rangle$ for the value of the functional μ on the test-function f.

Definition 2.12. For a symmetric distribution (generalized function) μ on $\mathbb{R}^N$, its *Bessel generating function* (or BGF) is a function of $(x_1, \ldots, x_N)$ given by

$$G_\theta(x_1, \ldots, x_N; \mu) := \frac{1}{N!}\left\langle \mu, B_{(a_1,\ldots,a_N)}(x_1, \ldots, x_N; \theta)\right\rangle, \tag{2.11}$$

where in the right-hand side $B_{(a_1,\ldots,a_N)}(x_1, \ldots, x_N; \theta)$ is treated as a test-function in $(a_1, \ldots, a_N)$ variables with parameters $(x_1, \ldots, x_N)$.

There are two tricky points in this definition. First, the N-tuple $(a_1, \ldots, a_N)$ was ordered in the original definition of the multivariate Bessel function, whereas μ is a distribution on $\mathbb{R}^N$. However, multivariate Bessel functions can be extended to $\mathbb{R}^N$ in a symmetric way. The $\frac{1}{N!}$ prefactor is introduced to match the integral over *ordered* N-tuples in (2.8) with distribution on whole $\mathbb{R}^N$ in (2.11).

More importantly, for general distributions μ (2.11) is not defined, since $B_{(a_1,\ldots,a_N)}(x_1, \ldots, x_N; \theta)$ is not compactly supported and therefore not a valid test function. Hence, one needs to impose some growth conditions similar to Definition 2.8 on μ, in order to make (2.11) meaningful. Rather than exploring the full generality, let us only consider the case of compactly supported μ (which means that μ vanishes on any test function whose support does not intersect a certain compact set), which is all we need for our application. For compactly supported distributions μ, Definition 2.12 is well-posed and in fact the pairing $\langle \mu, f\rangle$ makes sense for any infinitely-differentiable function f. In this text, we will be interested in compactly supported distributions μ of total mass equal to $N!$, meaning that $\langle \mu, \mathbf{1}\rangle = N!$, where $\mathbf{1}$ is the test function on $\mathbb{R}^N$ that is identically equal to 1; in this case, $G_\theta(0, \ldots, 0; \mu) = 1$.

Proposition 2.13. *Suppose that μ is a symmetric compactly supported distribution on $\mathbb{R}^N$. Then its BGF $G_\theta(x_1, \ldots, x_N; \mu)$ is an entire function. We also have*

$$\left(\prod_{i=1}^{s} \mathcal{P}_{k_i}\right) G_\theta(x_1, \ldots, x_N; \mu)\Big|_{x_1=\cdots=x_N=0} = \frac{1}{N!}\left\langle \mu, \prod_{i=1}^{s}\left(\sum_{j=1}^{N}(a_j)^{k_i}\right)\right\rangle. \tag{2.12}$$

Proof. Each compactly supported distribution can be identified with a (higher order) derivative of a compactly supported continuous function (see, e.g., [R91, Section 6]). Hence, we have

$$G_\theta(x_1, \ldots, x_N; \mu) = \int_{\mathbb{R}^N} \frac{\partial^{|\alpha|}}{\partial (a_i)^\alpha}\left[B_{(a_1,\ldots,a_N)}(x_1, \ldots, x_N; \theta)\right] f(a_1, \ldots, a_N)\, \mathrm{d}a_1 \cdots \mathrm{d}a_N,$$

[7] The space of test-functions f is equipped with a topology: f^n converge to 0 as $n \to \infty$, if the supports of all these functions belong to the same compact set and all partial derivatives of f^n converge to 0 uniformly.

where f is a compactly supported continuous function and $\frac{\partial^{|\alpha|}}{\partial(a_i)^\alpha}$ is a partial derivative of multi-index α in variables $(a_1, \dots, a_N)$. It remains to repeat the arguments of Sect. 2.3. We remark that the condition of μ being exponentially decaying, required by Proposition 2.11, has been substituted by the condition of μ being compactly supported. □

3. Statements of the Main Results

Throughout this section, we fix a real parameter $\gamma > 0$.

3.1. Law of large numbers at high temperature. Let $\{\mu_N\}_{N\geq 1}$ be a sequence of exponentially decaying probability measures, such that $\mu_N \in \mathcal{M}_N$ for each N, that is, μ_N is a probability measure on N-tuples $a_1 \leq a_2 \leq \cdots \leq a_N$. Alternatively, we can assume that each μ_N is a compactly supported symmetric distribution on $\mathbb{R}^N$ of total mass (i.e., the pairing against test function 1) equal to $N!$. Denote their Bessel generating functions by

$$G_{N;\theta}(x_1, \dots, x_N) := G_\theta(x_1, \dots, x_N; \mu_N).$$

By the results from the previous section, each $G_{N;\theta}(x_1, \dots, x_N)$ is holomorphic in a neighborhood of the origin and satisfies $G_{N;\theta}(0, \dots, 0) = 1$. Thus, for each N, the logarithm $\ln(G_{N;\theta})$ is a well-defined holomorphic function in a neighborhood of $(0, \dots, 0) \in \mathbb{C}^N$, and

$$\ln(G_{N;\theta})\big|_{x_1=\cdots=x_N=0} = 0.$$

We are interested in the interplay between the partial derivatives of $\ln(G_{N;\theta})$ at the origin and asymptotic properties of random μ_N-distributed[8] N-tuples $(a_1, \dots, a_N)$. We deal with the latter through the random variables

$$p_k^N := \frac{1}{N}\sum_{i=1}^{N}(a_i)^k, \qquad (a_1, \dots, a_N) \text{ is } \mu_N - \text{distributed}.$$

Definition 3.1 *(LLN-satisfaction).* We say that a sequence $\{\mu_N\}_{N\geq 1}$ *satisfies a Law of Large Numbers* if there exist real numbers $\{m_k\}_{k\geq 1}$ such that for any $s = 1, 2, \dots$ and any $k_1, \dots, k_s \in \mathbb{Z}_{\geq 1}$, we have

$$\lim_{N\to\infty} \mathbb{E}_{\mu_N} \prod_{i=1}^{s} p_{k_i}^N = \prod_{i=1}^{s} m_{k_i}.$$

Remark 3.2. Consider the empirical measure of $(a_1, \dots, a_N)$ given by $\frac{1}{N}\sum_{i=1}^{N}\delta_{a_i}$, where δ_x is the Dirac delta mass at $x \in \mathbb{R}$. Since the N-tuples $(a_1, \dots, a_N)$ are random, their empirical measures are random probability measures on $\mathbb{R}$. Under mild technical conditions (uniqueness of a solution to the moments problem, which holds whenever the numbers m_k do not grow too fast, see, e.g., [F71, Section VII.3]), LLN-satisfaction implies that these measures converge weakly, in probability, to a non-random measure whose moments are $m_1, m_2, \dots$.

[8] In our wordings we stick to the situation when μ_N are bona fide probability measures. If they are distributions (i.e. generalized functions possibly without any positivity), then all the random variables produced from them should be interpreted in formal sense: the laws of such random variables can be identified with expectations of various smooth functions of them, which are readily computed as pairings of μ_N with appropriate test functions. (One also should divide by $N!$ to adjust for differences between ordered and arbitrary N-tuples).

Definition 3.3 (γ*-LLN-appropriateness*). We say that the sequence $\{\mu_N\}_{N\geq 1}$ is γ*-LLN-appropriate* if there exists a sequence of real numbers $\{\kappa_l\}_{l\geq 1}$ such that

(a) $$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \frac{\partial^l}{\partial x_i^l} \ln\left(G_{N;\theta}\right)\Big|_{x_1=\cdots=x_N=0} = (l-1)!\cdot\kappa_l, \quad \text{for all } l, i \in \mathbb{Z}_{\geq 1}.$$

(b) $$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \frac{\partial}{\partial x_{i_1}}\cdots\frac{\partial}{\partial x_{i_r}} \ln\left(G_{N;\theta}\right)\Big|_{x_1=\cdots=x_N=0} = 0, \quad \text{for all } r\geq 2, \text{and } i_1,\ldots,i_r \in$$
$\mathbb{Z}_{\geq 1}$ such that the set $\{i_1,\ldots,i_r\}$ is of cardinality at least two.

Remark 3.4. Because the BGF $G_{N;\theta}(x_1,\ldots,x_N)$ is symmetric on the variables $x_1,\ldots,x_N$, the condition (a) is equivalent to:

(a') $$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \frac{\partial^l}{\partial x_1^l} \ln\left(G_{N;\theta}\right)\Big|_{x_1=\cdots=x_N=0} = (l-1)!\cdot\kappa_l, \quad \text{for all } l \in \mathbb{Z}_{\geq 1}.$$

Likewise, we could also simplify condition (b).

Remark 3.5. Suppose that

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \frac{\partial}{\partial z} \ln(G_{N;\theta}(z,0,\ldots,0)) = g(z),$$

uniformly over z in a complex neighborhood of 0. Then κ_l are the Taylor coefficients of g, that is,

$$g(z) = \sum_{l=1}^{\infty} \kappa_l z^{l-1}.$$

To state the main theorem, we use the language of formal power series in a formal variable z, namely series of the form

$$h_0 + h_1 z + h_2 z^2 + \cdots.$$

Definition 3.6. Let $\mathbb{R}[[z]]$ be the space of formal power series in z with real coefficients. Let $a(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ be any power series in $\mathbb{R}[[z]]$. We define three operators in $\mathbb{R}[[z]]$ by their action on a generic element $h(z) = h_0 + h_1 z + h_2 z^2 + \cdots \in \mathbb{R}[[z]]$, as follows.

- Derivation operator ∂:

$$\partial(h_0 + h_1 z + h_2 z^2 + \cdots) := h_1 + 2h_2 z + 3h_3 z^2 + \cdots.$$

- Lowering operator d:

$$d(h_0 + h_1 z + h_2 z^2 + \cdots) := h_1 + h_2 z + h_3 z^2 + \cdots.$$

- Multiplication operator $*_a$:

$$*_a(h(z)) := h(z)a(z).$$

Definition 3.7. Define the map $T^{\gamma}_{\kappa\to m} : \mathbb{R}^{\infty} \to \mathbb{R}^{\infty}$ that takes as input a countable real sequence $\{\kappa_l\}_{l\geq 1}$ and outputs the countable real sequence $\{m_k\}_{k\geq 1}$ by means of the relations

$$m_k = [z^0](\partial + \gamma d + *_g)^{k-1}(g(z)), \quad k = 1, 2, \ldots, \tag{3.1}$$

where $[z^0]$ is the constant term of the expression following it and

$$g(z) = \sum_{l=1}^{\infty} \kappa_l z^{l-1}.$$

For notation purposes, in the remainder of the paper the input of the map $T^{\gamma}_{\kappa\to m}$ is denoted by $\{\kappa_l\}_{l\geq 1}$ and the output is denoted by $\{m_k\}_{k\geq 1}$. Whenever $T^{\gamma}_{\kappa\to m}(\{\kappa_l\}_{l\geq 1}) = \{m_k\}_{k\geq 1}$, the quantities κ_l are called *γ-cumulants* and the m_k's are called *moments*. This is meant to draw an analogy with the sequences of classical cumulants and moments of a probability measure. The motivation for this terminology is explained by the results in Sect. 8. Roughly speaking, the map $T^{\gamma}_{\kappa\to m}$ degenerates to the relation between cumulants and moments when $\gamma \to 0$, and to the relation between free cumulants and moments when $\gamma \to \infty$.

Theorem 3.8 (Law of Large Numbers for high temperature). *The sequence $\{\mu_N\}_{N\geq 1}$ is γ-LLN-appropriate if and only if it satisfies a LLN. In case this occurs, the sequences $\{\kappa_l\}_{l\geq 1}$ and $\{m_k\}_{k\geq 1}$ are related by*

$$\{m_k\}_{k\geq 1} = T^{\gamma}_{\kappa\to m}(\{\kappa_l\}_{l\geq 1}). \tag{3.2}$$

The proof of this theorem is given later in Sect. 5 below.

Our next results describe in more detail the map $T^{\gamma}_{\kappa\to m}$ from Definition 3.7.

3.2. Combinatorial formula for the map $T^{\gamma}_{\kappa\to m}$. From Definition 3.7, we are able to obtain the values of m_k by doing calculations with formal power series and isolating the constant term of the resulting expansion. For example, for $k = 1, 2, 3, 4$, the resulting formulas are the following:

$$\begin{aligned}
m_1 &= \kappa_1,\\
m_2 &= (\gamma+1)\kappa_2 + \kappa_1^2,\\
m_3 &= (\gamma+1)(\gamma+2)\kappa_3 + 3(\gamma+1)\kappa_2\kappa_1 + \kappa_1^3,\\
m_4 &= (\gamma+1)(\gamma+2)(\gamma+3)\kappa_4 + 4(\gamma+1)(\gamma+2)\kappa_3\kappa_1\\
&\quad + (\gamma+1)(2\gamma+3)\kappa_2^2 + 6(\gamma+1)\kappa_2\kappa_1^2 + \kappa_1^4.
\end{aligned} \tag{3.3}$$

However, the defining formula (3.1) is not explicit enough and becomes complicated when k is large. Our next main theorem is a simpler combinatorial formula that expresses m_k as a polynomial of the variables $\kappa_1, \kappa_2, \ldots, \kappa_k$. To state it, we need some terminology.

For any $k \in \mathbb{Z}_{\geq 1}$, denote $[k] := \{1, 2, \ldots, k\}$. A *set partition* π of $[k]$ is an (unordered) collection of pairwise disjoint nonempty subsets of $[k]$ such that $[k] = B_1 \cup \cdots \cup B_m$. The subsets $B_1, \ldots, B_m$ are called the *blocks* of the set partition π and we use the notation $\pi = B_1 \sqcup \cdots \sqcup B_m$. The cardinalities of the blocks are denoted $|B_1|, \ldots, |B_m|$. We denote the collection of all set partitions of $[k]$ by $\mathscr{P}(k)$. Given a set partition π, we denote by $\#(\pi)$ its number of blocks. For example, there are seven set partitions of [4] with two blocks; they are:

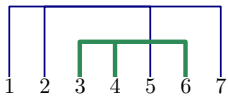

Fig. 1. Set partition $\{1, 2, 5, 7\} \sqcup \{3, 4, 6\}$

$$\{1\} \sqcup \{2, 3, 4\}, \quad \{1, 3, 4\} \sqcup \{2\}, \quad \{1, 2, 4\} \sqcup \{3\}, \quad \{1, 2, 3\} \sqcup \{4\},$$
$$\{1, 2\} \sqcup \{3, 4\}, \quad \{1, 3\} \sqcup \{2, 4\}, \quad \{1, 4\} \sqcup \{2, 3\}.$$

We also use the Pochhammer symbol notation:

$$(x)_q := \begin{cases} x(x+1)\cdots(x+q-1), & \text{if } q \in \mathbb{Z}_{\geq 1}, \\ 1, & \text{if } q = 0. \end{cases}$$

Definition 3.9. For any $\pi \in \mathscr{P}(k)$ and $\gamma \in \mathbb{R}$, define the quantity $W(\pi)$, that will be called the *γ-weight of π*, as follows[9]. Let $m = \#(\pi)$ and label the blocks of π by $B_1, \ldots, B_m$ in such a way that the smallest element from B_i is smaller than all elements from B_j, whenever $i < j$. That is, if the blocks are $B_i = \{b_1^i < \cdots < b_{|B_i|}^i\}$, then $b_1^1 < b_1^2 < \cdots < b_1^m$. For each $i \in \{1, \ldots, m\}$, define $p(i)$ as the number of indices $j \in \{1, \ldots, |B_i| - 1\}$ such that $\{b_j^i + 1, \ldots, b_{j+1}^i - 1\} \cap B_t \neq \emptyset$ for some block B_t with $t < i$, and set $q(i) := |B_i| - 1 - p(i)$. Then define

$$W(\pi) := \prod_{i=1}^{m} \Big(p(i)! \cdot (\gamma + p(i) + 1)_{q(i)} \Big). \tag{3.4}$$

For a set partition $\pi = B_1 \sqcup \cdots \sqcup B_m$ of $[k]$, we can think of the quantity $p(i)! \cdot (\gamma + p(i) + 1)_{q(i)}$ as a weight associated to the block B_i. Therefore the γ-weight $W(\pi)$ is the product of all weights of the blocks of π. The weight of a block B_i depends on the integer $p(i)$, whose computation can be visualized through a geometric procedure involving arc diagrams:

- Draw each block $B_j = (b_1^j < b_2^j < \cdots < b_r^j)$ as an arc with r vertical legs at positions b_a^j, $a = 1, \ldots, r$ and with $r - 1$ horizontal roofs joining adjacent legs at height $m + 1 - j$.
- $p(i)$ is the number of roofs in B_i, which intersect legs of other blocks. Note that each roof is counted only once, no matter how many legs it intersects.

Let us provide several examples. First, consider set partition $\{1, 2, 5, 7\} \sqcup \{3, 4, 6\} \in \mathscr{P}(7)$ corresponding to the following arc diagram:
The blocks are labeled $B_1 = \{1, 2, 5, 7\}$ and $B_2 = \{3, 4, 6\}$. We have $p(1) = 0$, $q(1) = 3$, $p(2) = 1$, $q(2) = 1$, and therefore

$$W(\pi) = 0! \cdot (\gamma + 1)_3 \cdot 1! \cdot (\gamma + 2)_1 = (\gamma + 1)(\gamma + 2)^2(\gamma + 3).$$

For a different example, consider set partition $\{1, 4\} \sqcup \{2, 6\} \sqcup \{3, 5, 7\} \in \mathscr{P}(7)$ corresponding to the following arc diagram:

[9] We omit the dependence on γ from the notation $W(\pi)$.

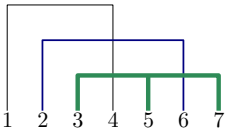

Fig. 2. Set partition $\{1, 4\} \sqcup \{2, 6\} \sqcup \{3, 5, 7\}$

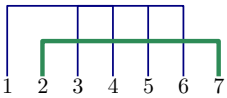

Fig. 3. Set partition $\{1, 3, 4, 5, 6\} \sqcup \{2, 7\}$

The blocks are labeled $B_1 = \{1, 4\}$, $B_2 = \{2, 6\}$, and $B_3 = \{3, 5, 7\}$. We have $p(1) = 0$, $q(1) = 1$, $p(2) = 1$, $q(2) = 0$, $p(3) = 2$, $q(3) = 0$, and therefore

$$W(\pi) = 0! \cdot (\gamma + 1)_1 \cdot 1! \cdot (\gamma + 2)_0 \cdot 2! \cdot (\gamma + 3)_0 = 2(\gamma + 1).$$

For the final example, consider set partition $\{1, 3, 4, 5, 6\} \sqcup \{2, 7\} \in \mathscr{P}(7)$ corresponding to the following arc diagram:
The blocks are labeled $B_1 = \{1, 3, 4, 5, 6\}$ and $B_2 = \{2, 7\}$. We have $p(1) = 0$, $q(1) = 4$, $p(2) = 1$, $q(2) = 0$, and therefore

$$W(\pi) = 0! \cdot (\gamma + 1)_4 \cdot 1! \cdot (\gamma + 1)_0 = (\gamma + 1)(\gamma + 2)(\gamma + 3)(\gamma + 4).$$

Let us also mention two useful properties which directly follow from the definition of $p(i)$:

- $p(1) = 0$.
- If $|B_i| = 1$, then $p(i) = q(i) = 0$.

We have introduced all notations and can now state the main theorem of this section:

Theorem 3.10 (γ-cumulants to moments formula). *Let $\{m_k\}_{k\geq 1}$ and $\{\kappa_l\}_{l\geq 1}$ be real sequences that are related by $\{m_k\}_{k\geq 1} = T^{\gamma}_{\kappa\to m}(\{\kappa_l\}_{l\geq 1})$. Let $k \in \mathbb{Z}_{\geq 1}$ be arbitrary. Then*

$$m_k = \sum_{\pi\in\mathscr{P}(k)} W(\pi) \prod_{B\in\pi} \kappa_{|B|}.$$

The proof is presented in Sect. 6 below. In Sect. 8 we explain how in the limits $\gamma \to 0$ and $\gamma \to \infty$, Theorem 3.10 turns into the expression of moments through classical cumulants and through free cumulants, respectively.

3.3. $T^{\gamma}_{\kappa\to m}$ and its inverse $T^{\gamma}_{m\to\kappa}$ through generating functions. The map $T^{\gamma}_{\kappa\to m} : \{\kappa_l\}_{l\geq 1} \mapsto \{m_k\}_{k\geq 1}$ is equivalent to relations of the form

$$m_k = (\gamma + 1)_{k-1} \cdot \kappa_k + \text{ certain polynomial in the variables } \kappa_1, \ldots, \kappa_{k-1}, \qquad (3.5)$$

Recursively using (3.5), each κ_l can be expressed as a polynomial in the variables $m_1, \ldots, m_l$. In other words, the map $T^{\gamma}_{\kappa\to m}$ has an inverse denoted by

$$T^{\gamma}_{m\to\kappa} := (T^{\gamma}_{\kappa\to m})^{-1} : \{m_k\}_{k\geq 1} \mapsto \{\kappa_l\}_{l\geq 1}.$$

For example, inverting the formulas in (3.3) we get

$$\begin{aligned}
\kappa_1 &= m_1,\\
\kappa_2 &= \frac{1}{\gamma+1}\left(m_2 - m_1^2\right),\\
\kappa_3 &= \frac{1}{(\gamma+1)_2}\left(m_3 - 3m_2m_1 + 2m_1^3\right),\\
\kappa_4 &= \frac{1}{(\gamma+1)_3}\left(m_4 - 4m_3m_1 - \left[2+\frac{1}{\gamma+1}\right]m_2^2 + \left[10+\frac{2}{\gamma+1}\right]m_2m_1^2 - \left[5+\frac{1}{\gamma+1}\right]m_1^4\right).
\end{aligned} \tag{3.6}$$

One way to write the formulas connecting moments and cumulants in a compact form is through generating function:

Theorem 3.11. *Let* $\{m_k\}_{k\geq 1}$ *and* $\{\kappa_l\}_{l\geq 1}$ *be real sequences related by* $\{\kappa_l\}_{l\geq 1} = T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\geq 1})$. *Then*

$$\exp\left(\sum_{l=1}^{\infty}\frac{\kappa_l y^l}{l}\right) = [z^0]\left\{\sum_{n=0}^{\infty}\frac{(yz)^n}{(\gamma)_n}\cdot\exp\left(\gamma\sum_{k=1}^{\infty}\frac{m_k}{k}z^{-k}\right)\right\}. \tag{3.7}$$

Equivalently, (3.7) *can be rewritten as a combination of two identities involving an auxiliary sequence* $\{c_n\}_{n\geq 0}$ *through:*

$$\begin{cases}
\exp\left(\sum_{l=1}^{\infty}\frac{\kappa_l}{l}z^l\right) = \sum_{n=0}^{\infty}\frac{c_n}{(\gamma)_n}z^n,\\
\exp\left(\gamma\sum_{k=1}^{\infty}\frac{m_k z^k}{k}\right) = \sum_{n=0}^{\infty}c_n z^n.
\end{cases} \tag{3.8}$$

As we explain in Sect. 8, in the limit $\gamma \to 0$, the statement of Theorem 3.11 turns into the well-known formula expressing the generating function of (classical) cumulants as a logarithm of the generating function of moments (equivalently, of the characteristic function of a random variable). On the other hand, in the limit $\gamma \to \infty$, Theorem 3.11 can be converted into the identification of the free cumulants with Taylor series coefficients of the Voiculescu R-transform of a probability measure.

A close examination of (3.8) reveals an unexpected connection to the d-cumulants for the (additive) finite free convolution. We recall that the latter is a deterministic binary operation on d-tuples of real numbers, which was shown in [GM20, Theorem 1.2] to be the $\theta \to \infty$ limit of the operation $(\mathbf{a}, \mathbf{b}) \mapsto \mathbf{a} +_\theta \mathbf{b}$ for fixed $N = d$. Generating functions and certain combinatorial formulas for d-cumulants were developed in [Ma18, AP18]. Comparing with [AP18], we observe a match under the following change in notations, where in the left column we use notations from [AP18] and in the right column we use notations from our work:

$$\begin{aligned}
d &\longleftrightarrow -\gamma,\\
m_n &\longleftrightarrow m_n,\\
\kappa_n &\longleftrightarrow \gamma^{n-1}\kappa_n,\\
a_n &\longleftrightarrow (-1)^n c_n.
\end{aligned} \tag{3.9}$$

Indeed, under (3.9) the first formula of (3.8) becomes [AP18, (3.1) or (3.3)] and the second formula of (3.8) becomes [AP18, (4.2)]. Note that the symbol $(x)_n$ has the meaning $x(x-1)\ldots(x-n+1)$ in [AP18], which is different from the convention that we use.

It is important to emphasize that in our work $\gamma > 0$, while in [Ma18, AP18], d is a positive integer. Hence, using (3.9) we see that there are no values of parameters under which finite free cumulants coincide with our γ-cumulants. Instead, one family of cumulants should be treated as an analytic continuation of another. There are two consequences of this correspondence. First, Theorem 3.10 translates into a new combinatorial formula for finite free cumulants. Second, [AP18, Theorem 4.2] explains how the generating function identity equivalent to (3.8) leads to transition formulas (involving double sums over set partitions) between moments and finite free cumulants and vice versa. Hence, substituting (3.9) we can obtain similar formulas between moments and our γ-cumulants.

3.4. Generalized Markov–Krein transform. There is a way to recast the formulas of Theorem 3.11 connecting them to a remarkable non-linear transformation of measures discussed in [FF16] (see also [MP21] and [K98]). We take the numbers c_n from (3.8) and replace them with

$$\tilde{c}_n = \frac{n!}{(\gamma)_n} c_n.$$

Further, suppose that m_k are moments of a compactly supported probability measure ν and $\tilde{c}_n$ are moments of a compactly supported probability measure μ:

$$m_k = \int_{\mathbb{R}} x^k \nu(dx), \qquad \tilde{c}_n = \int_{\mathbb{R}} x^n \mu(dx).$$

Then the second identity of (3.8) can be recast as

$$\exp\left(-\gamma \int_{\mathbb{R}} \ln(z-x)\nu(dx)\right) = \int_{\mathbb{R}} \frac{1}{(z-t)^\gamma} \mu(dt), \tag{3.10}$$

where the equivalence of (3.10) with (3.8) can be seen by assuming z to be large and expanding the integrals into $1/z$ power series. It is proven in [FF16] that for any probability measure ν with $\int_{\mathbb{R}} \ln(1+|x|)\nu(dx) < \infty$ (in particular, compact support is not necessary), there exists another probability measure μ, such that the identity (3.10) holds. For $\gamma = 1$ the correspondence (3.10) and its relatives were popularized in the context of asymptotic problems by Kerov (see [K98], [K03, Chapter VI]) under the name Markov–Krein transform; its origins go back to the studies of the solutions to the moment problems in the middle of the XX century. General $\gamma > 0$ case was mentioned in [K98, Section 3.7 and 4.1] and further discussed in [FF16]. In our setting the correspondence (3.8) is useful because the first identity of (3.8) is recast in terms of the measure μ as

$$\sum_{l=1}^{\infty} \frac{\kappa_l}{l} z^l = \ln\left(\int_{\mathbb{R}} \exp(tz)\mu(dt)\right).$$

Therefore, up to multiplication by $(l-1)!$, the γ-cumulants κ_l of the measure ν are classical cumulants of the measure μ (we recall the definition of the classical cumulants in Sect. 8.1).

In a sense, the correspondence $\nu \leftrightarrow \mu$ of (3.10) reduces γ-cumulants (and all operations based on them) to classical cumulants. However, a difficulty in efficiently using this point of view is that the correspondence is highly non-linear and its properties are mostly unknown. For instance, describing all measures μ, which can appear in the right-hand side of (3.10) is an open question.

4. Applications

In this section we list several applications of the general theorems from Sect. 3.

4.1. Law of large numbers for Gaussian β ensembles. For each $N \geq 1$, let $\mu_{N,\theta}$ be the N-particle *Gaussian β ensemble* with parameter $\beta = 2\theta$—this is a probability distribution on N-tuples of real numbers $a_1 \leq a_2 \leq \cdots \leq a_N$ with density proportional to

$$\prod_{1 \leq i < j \leq N} (a_j - a_i)^{2\theta} \prod_{i=1}^{N} e^{-a_i^2/2}. \tag{4.1}$$

The eigenvalue distributions of the celebrated Gaussian Orthogonal/Unitary/Symplectic ensembles of random matrices are given by (4.1) at $\theta = \frac{1}{2}/1/2$, respectively.

To state the result of this subsection, we need a few definitions. Denote by $\mathscr{M}(k)$ the collection of all *perfect matchings* of $[k]$, that is, the collection of set partitions of $[k]$ where each block has size 2. $\mathscr{M}(k)$ is empty if k is odd, and if $k = 2m$ is even, then $\mathscr{M}(k)$ has cardinality $(2m-1)!! = (2m-1)(2m-3)\cdots 3 \cdot 1$. Any perfect matching $\pi = \{B_1, \ldots, B_m\}$ of $[2m]$ is also a set partition of $[2m]$, so we can draw its arc diagram, as described in Sect. 3.2. Denote by $\mathrm{roof}(\pi)$ the number of roofs that do not intersect some leg. $\mathrm{Roof}(\pi)$ is an integer between 1 and m, and $\mathrm{roof}(\pi) = m$ if and only if the perfect matching π is non-crossing, see Fig. 4 for an illustration.

Finally, consider the empirical measures

$$\rho_{N,\theta} := \frac{1}{N} \sum_{i=1}^{N} \delta_{a_i}, \qquad (a_1, \ldots, a_N) \text{ is } \mu_{N,\theta}\text{-distributed.}$$

Theorem 4.1. *As $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$, the (random) measures $\rho_{N,\theta}$ converge weakly, in probability to a deterministic probability measure μ_γ which is uniquely determined by its moments:*

$$\int_{-\infty}^{\infty} x^k \mu_\gamma(dx) = \sum_{\pi \in \mathscr{M}(k)} (\gamma + 1)^{\mathrm{roof}(\pi)}, \tag{4.2}$$

which is set to be 0 *for odd k.*

Remark 4.2. In our Theorem 4.1, the limiting measure μ_γ is an analogue of Wigner's semicircle law from free probability theory and of the Gaussian distribution from classical probability, because the only nonzero γ-cumulant is the second one. Similarly to these measures, μ_γ is also present in a Central Limit Theorem with respect to the operation of γ-convolution discussed in the next subsection, see [MP21, Section 5.3]. In fact, μ_γ degenerates into these measures at special values of γ. Indeed, when $\gamma = 0$ and $k = 2m$, the right-hand side of (4.2) is equal to $|\mathscr{M}(2m)| = (2m-1)!!$, which coincides with the $(2m)$-th moment of the standard normal distribution. When $\gamma \to \infty$ and $k = 2m$, the right-hand side of (4.2) (when divided by γ^m) becomes the number of non-crossing perfect matchings of $[2m]$, which is the m-th Catalan number $C_m = \frac{1}{m+1}\binom{2m}{m}$. This is the $(2m)$-th moment of the standard Wigner's semicircle law.

Remark 4.3. While identification of μ_γ through (4.2) was not stated explicitly in the literature before, the LLN itself, i.e. existence of the limiting measures μ_γ is known at least from [ABG12]. [BP15] provides other (more complicated) formulas for the moments of μ_γ, and in [DS15] the measure μ_γ is identified with the mean spectral measure of a certain random Jacobi matrix.

Remark 4.4. The measures μ_γ were also previously studied by other authors without knowing about their connections to Gaussian β ensembles. Askey and Wimp [AW84] studied μ_γ as an orthogonality measure for the *associated Hermite polynomials* ([D09] obtained a formula for the moments that is equivalent to ours, and [SZ20,BDEG21] contain generalizations: see (4.7) in the former paper and Proposition 4.19 in the latter). Interestingly, the same polynomials also play a role in studying $\beta \to \infty$ limits of Gaussian β Ensembles, see [GK20, Section 4.3]. From another direction, Kerov [K98] studied μ_γ in connection to the Markov-Krein transform and noticed that these measures interpolate between Gaussian and semicircle laws.

Proof of Theorem 4.1. By Lemma 4.5 below, the Bessel generating function of $\mu_{N,\theta}$ is

$$G_N(x_1, \dots, x_N; \theta) = \exp\left(\frac{x_1^2 + \dots + x_N^2}{2}\right).$$

It follows that $\{\mu_{N,\theta}\}$ is γ-LLN-appropriate with $\{\kappa_l\}_{l\geq 1}$ given by

$$\kappa_l = \begin{cases} 1, & \text{if } l = 2, \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

The corresponding sequence $\{m_k\}_{k\geq 1} = T^{\gamma}_{\kappa\to m}(\{\kappa_l\}_{l\geq 1})$ is given by the formula in Theorem 3.10. Because the only nonzero γ-cumulant κ_l is the one with $l = 2$, the summation for m_k reduces from all set partitions of $[k]$ to all perfect matchings of $[k]$. In particular, $m_k = 0$ if k is odd. In the case that k is even, say $k = 2m$, consider any perfect matching $\pi = \{B_1, \dots, B_m\}$ of $[2m]$; each block B_i has cardinality 2, so $p(i)$ is 1 if the roof of the arc B_i intersects some leg in the arc-diagram of π, and otherwise $p(i)$ is 0. As a result, the weight $W(\pi)$ in (3.4) is equal to $(\gamma+1)^{\mathrm{roof}(\pi)}$.

Then Theorem 3.8 shows that the sequence $\{\mu_{N,\theta}\}$ satisfies a LLN, and this proves the desired convergence in the statement of the theorem, see Remark 3.2. It remains to show that the right-hand sides of (4.2) are the moments of a *unique* probability measure. For this, we check the Carleman's condition: the moments problem for a sequence $\{\alpha_k\}_{k\geq 1}$ determines a unique probability measure if

$$\sum_{m=1}^{\infty} (\alpha_{2m})^{-1/(2m)} = +\infty. \tag{4.4}$$

Indeed, elementary bounds show

$$\begin{aligned} \alpha_{2m} &= \sum_{\pi\in\mathscr{M}(2m)} (\gamma+1)^{\mathrm{roof}(\pi)} \leq (\gamma+1)^m \cdot |\mathscr{M}(2m)| = (\gamma+1)^m \cdot (2m-1)!! \\ &\leq (\gamma+1)^m \cdot (2m)^m \end{aligned}$$

and so $(\alpha_{2m})^{-1/(2m)} \geq \mathrm{const} \cdot \frac{1}{\sqrt{m}}$, thus proving (4.4). □

Lemma 4.5. *The Bessel generating function of the N-particle Gaussian β ensemble with parameter $\beta = 2\theta$ is*

$$G_\theta(x_1, \dots, x_N; \mu_{N,\theta}) = \exp\left(\frac{x_1^2 + \cdots + x_N^2}{2}\right). \tag{4.5}$$

Identity (4.5) is a folklore and we do not claim any novelty. One can prove it by taking an appropriate limit of the Cauchy identity for Jack polynomials. Alternatively, the computation of the Bessel generating function for the Laguerre β Ensemble is equivalent to [Ne03, Eqn. (3.1)], and (4.5) is then obtained by a limit transition. As yet another approach, Gaussian β Ensembles can be identified with measures from [AN19, Thm. 1.13 and Sec. 6] with a single non-zero parameter γ_2, and for those, the Bessel generating function is computed in [AN19], see also [OV96, Cor. 2.6] for the $\theta = 1$ case. For completeness, let us sketch the first argument.

Sketch of the proof of Lemma 4.5. We rely on the theory of symmetric functions, as presented in [M98, Ch. I and VI], see also [S89]. We denote by Λ the real algebra of symmetric functions in infinitely many variables $x_1, x_2, \dots$, which is generated by the (algebraically independent) power sums $p_k = \sum_i (x_i)^k$, $k = 1, 2, \dots$. A distinguished linear basis of Λ is given by the Jack symmetric functions $P_\lambda(\cdot; \theta)$, where $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq 0)$ ranges over the set of all partitions and $P_\lambda(\cdot; \theta)$ is homogeneous of degree $|\lambda| = \sum_i \lambda_i$. The Jack symmetric functions exhibit the following Cauchy(-Littlewood) summation identity:

$$\sum_\lambda \frac{P_\lambda(x_1, x_2, \dots; \theta) P_\lambda(y_1, y_2, \dots; \theta)}{\langle P_\lambda, P_\lambda \rangle} = \exp\left(\theta \sum_{k=1}^\infty \frac{p_k(x_1, x_2, \dots) p_k(y_1, y_2, \dots)}{k}\right), \tag{4.6}$$

where the sum is over all partitions λ and $\langle P_\lambda, P_\lambda \rangle$ are explicitly known θ-dependent normalization constants. (4.6) is a formal identity of power series. We can turn it into a numeric identity by applying a *specialization*: an algebra homomorphism from Λ to the real numbers. We need one particular homomorphism, which is known as the Plancherel specialization in the literature:

$$\tau_s : \Lambda \to \mathbb{R}, \qquad \tau_k(p_k) = \begin{cases} s, & k = 1, \\ 0, & k \geq 1. \end{cases}$$

Applying τ_s with $s > 0$ to the x-variables and setting $y_{N+1} = y_{N+2} = \cdots = 0$, we transform(4.6) into

$$\sum_{\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0)} \frac{P_\lambda(\tau_s; \theta) P_\lambda(1^N; \theta)}{\langle P_\lambda, P_\lambda \rangle} \cdot \frac{P_\lambda(y_1, \dots, y_N; \theta)}{P_\lambda(1^N; \theta)} = \exp\left(\theta s \sum_{i=1}^N y_i\right), \tag{4.7}$$

where 1^N stays for the N variables equal to 1 and the summation got restricted to the partitions with (at most) N parts, because $P_\lambda(y_1, \dots, y_N; \theta)$ vanishes for others. Let us now introduce a probability measure on N-tuples of integers $\lambda = (\lambda_1 \geq \dots \lambda_N \geq 0)$ through the formula:

$$\mathrm{Prob}(\lambda_1, \dots, \lambda_N) = \mathcal{J}_{1^N; \tau_s}(\lambda) := e^{-\theta s N} \frac{P_\lambda(\tau_s; \theta) P_\lambda(1^N; \theta)}{\langle P_\lambda, P_\lambda \rangle}. \tag{4.8}$$

$\mathcal{J}_{1^N;\tau_s}(\lambda)$ is an instance of the Jack measure on partitions and it is discussed in details in [GSh15, Sec. 2.2]. With this definition, we can rewrite (4.7) as

$$\mathbb{E}_{\mathcal{J}_{1^N;\tau_s}}\left[\frac{P_\lambda(y_1,\dots,y_N;\theta)}{P_\lambda(1^N;\theta)}\right]=\exp\left(\theta s\sum_{i=1}^N(y_i-1)\right). \tag{4.9}$$

Next, we want to set

$$s=\varepsilon^{-1}\theta^{-1},\qquad \lambda_i=\varepsilon^{-1}+\varepsilon^{-1/2}a_{N+1-i},\quad y_i=\exp(\varepsilon^{1/2}x_i),\quad i=1,2,\dots,N, \tag{4.10}$$

for $\varepsilon>0$ and then send $\varepsilon\to 0$. [GSh15, Proposition 2.10 with $t=1$] proves that the measures (4.8) converge weakly to the Gaussian β ensemble of (4.1):

$$\lim_{\varepsilon\to 0+}\mathcal{J}_{1^N;\,\tau_{\varepsilon^{-1}\theta^{-1}}}=\mu_{N,\theta}.$$

In this way (4.5) is obtained as $\varepsilon\to 0$ limit of (4.9). For the right-hand side of (4.9) we have:

$$\exp\left(\theta s\sum_{i=1}^N(y_i-1)\right)=\exp\left(\varepsilon^{-1/2}\sum_{i=1}^N x_i+\sum_{i=1}^N\frac{x_i^2}{2}+O(\varepsilon^{1/2})\right). \tag{4.11}$$

For the left hand side of (4.7), we use the identity (valid for any $c\in\mathbb{Z}$) which is a discrete version of (1.8):

$$P_{(c+\mu_1,\dots,\,c+\mu_N)}(y_1,\dots,y_N)=(y_1\cdots y_N)^c\cdot P_{(\mu_1,\dots,\,\mu_N)}(y_1,\dots,y_N)$$

and limit relation between the symmetric Jack polynomials and multivariate Bessel functions:

$$\lim_{\eta\to 0^+}\frac{P_{(\eta^{-1}a_N,\dots,\eta^{-1}a_1)}(\exp(\eta x_1),\dots,\exp(\eta x_N);\theta)}{P_{(\eta^{-1}a_N,\dots,\eta^{-1}a_1)}(1^N;\theta)}=B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\theta). \tag{4.12}$$

The last identity can be found in [OO97, Sec. 4] or [Cu19, Thm. 7.5]; in this limit transition the combinatorial formula for Jack polynomials (expressing the expansion of the polynomial into monomials as a sum over semi-standard Young tableaux or Gelfand-Tsetlin patterns) turns into (2.5) of Definition 2.5.

Thus, dividing both sides of (4.7) by $\exp\left(\varepsilon^{-1/2}\sum_{i=1}^N x_i\right)$ and sending $\varepsilon\to 0$ (we omit a standard tail bound justifying the validity of the limit transition under the expectation sign), we get

$$\mathbb{E}_{\mu_{N,\theta}}\left[B_{(a_1,\dots,a_N)}(x_1,\dots,x_N;\theta)\right]=\exp\left(\sum_{i=1}^N\frac{x_i^2}{2}\right).$$

□

4.2. γ-convolution.

Proof of Theorem 1.5. Let $G_N^{\mathbf{a}}$ be the BGF of the distribution[10] of $\mathbf{a}(N)$ and let $G_N^{\mathbf{b}}$ be the BGF of the distribution of $\mathbf{b}(N)$. Since Definition 1.4 is the same as Definition 3.1, Theorem 3.8 implies that the distributions of $\mathbf{a}(N)$ and $\mathbf{b}(N)$ are γ-LLN appropriate. Let us denote the corresponding γ-cumulants (right-hand sides in (a) of Definition 3.3) through $\kappa_l^{\mathbf{a}}$ and $\kappa_l^{\mathbf{b}}$, respectively.

Further, let $G_N^{\mathbf{a}+_\theta\mathbf{b}}$ be the BGF of $\mathbf{a}(N)+_\theta\mathbf{b}(N)$. By Definition 1.2 and the independence of the distributions of $\mathbf{a}(N)$ and $\mathbf{b}(N)$, we have

$$G_N^{\mathbf{a}+_\theta\mathbf{b}}(x_1,\dots,x_N;\,\theta)=G_N^{\mathbf{a}}(x_1,\dots,x_N;\,\theta)\cdot G_N^{\mathbf{b}}(x_1,\dots,x_N;\,\theta).$$

Hence, partial derivatives of $\ln(G_N^{\mathbf{a}+_\theta\mathbf{b}})$ are sums of those of $\ln(G_N^{\mathbf{a}})$ and those of $\ln(G_N^{\mathbf{a}})$. Therefore, the sequence of distributions of $\mathbf{a}(N)+_\theta\mathbf{b}(N)$ is γ-LLN appropriate with γ-cumulants given by the sums $\kappa_l^{\mathbf{a}}+\kappa_l^{\mathbf{b}}$, $l=1,2,\dots$. Applying Theorem 3.8 again, we conclude that $\mathbf{a}(N)+_\theta\mathbf{b}(N)$ converges in the sense of moments, which concludes the proof of the theorem. Observe that the lth γ-cumulant of the limit measure is $\kappa_l^{\mathbf{a}}+\kappa_l^{\mathbf{b}}$, for all $l=1,2,\dots$, which then verifies formula (1.11). □

One remark is in order. We never prove (or claim) that the $N\to\infty$ limit of empirical distributions of $\mathbf{a}(N)+_\theta\mathbf{b}(N)$ is given by a probability measure; we only show that the moments converge to some limiting values. Of course, if we knew that $\mathbf{a}(N)+_\theta\mathbf{b}(N)$ is a bona fide random N-tuple of integers (which is widely believed to be true, see Conjecture 1.3), then we could say that the deterministic limit of random empirical distributions is necessarily given by a probability measure. In this case, the binary operation $\boxplus_\gamma$ would turn into an operation on probability measures.

Definition 4.6. Let $\gamma>0$ be fixed. Let μ be a probability measure with finite moments $\{m_k^\mu\}_{k\ge1}$. If $\{\kappa_l^\mu\}_{l\ge1}:=T^\gamma_{\kappa\to m}(\{m_k^\mu\}_{k\ge1})$, then the quantities κ_l^μ are said to be the *γ-cumulants of μ*. Next, let ν be another probability measure with all γ-cumulants $\{\kappa_l^\nu\}_{l\ge1}$. If there exists a unique probability measure, to be denoted $\mu\boxplus_\gamma\nu$, such that all of its γ-cumulants are finite and given by

$$\kappa_l^{\mu\boxplus_\gamma\nu}=\kappa_l^\mu+\kappa_l^\nu,\ \text{for all } l\ge1,$$

then $\mu\boxplus_\gamma\nu$ is said to be the *γ-convolution* of μ and ν.

We note that if Conjecture 1.3 is true, then the words "If there exists a probability measure" can be removed from the above definition, because such a measure would always exists. However, one should still require the uniqueness because not every probability measure is uniquely determined by its γ-cumulants or moments; one needs to impose a slow growth condition on the moments to guarantee the uniqueness.

This definition agrees with the definition embedded in Theorem 1.5 in the following sense. If τ is a probability measure which is determined by its moments $\{m_k^\tau\}_{k\ge1}$, then $\tau=\mu\boxplus_\gamma\nu$ (in the sense of Definition 4.6) if and only if $\{m_k^\tau\}_{k\ge1}=\{m_k^\mu\}_{k\ge1}\boxplus_\gamma\{m_k^\nu\}_{k\ge1}$ (in the sense of Theorem 1.5).

[10] In this section the word "distribution" is used in probabilistic meaning, as in "distribution of a random variable", rather than in functional-analytic meaning, where a distribution is a synonym of a generalized function.

4.3. Examples of γ-convolutions.

Example 4.7. Let $\{m_k\}_{k\geq 1}$ be the sequence of moments of a probability measure μ on $\mathbb{R}$. Also consider the sequence $\{a^k\}_{k\geq 1}$ of powers of a real number $a \in \mathbb{R}$; evidently this is the sequence of moments of the Dirac delta mass at point a. Let $\{\widetilde{m}_k\}_{k\geq 1}$ be the sequence of moments of the conventional convolution $\delta_a * \mu$, in other words, if we set $m_0 := 1$ then

$$\widetilde{m}_k = \sum_{i=0}^{k} \binom{k}{i} a^i m_{k-i}, \qquad k = 1, 2, \dots. \tag{4.13}$$

For any $\gamma > 0$, we claim

$$\{\widetilde{m}_k\}_{k\geq 1} = \{a^k\}_{k\geq 1} \boxplus_\gamma \{m_k\}_{k\geq 1}.$$

This is equivalent to $\delta_a * \mu = \delta_a \boxplus_\gamma \mu$, i.e. it would mean that γ-convolution with a Dirac delta mass at a is identified with shift by a. Indeed, we can verify this claim by using Theorem 1.5. Note that for the constant sequences $\{\mathbf{a}(N) = (a, \dots, a)\}_{N\geq 1}$, we have $\lim_{N\to\infty} \mathbf{a}(N) \overset{m}{=} \{a^k\}_{k\geq 1}$. Moreover, for any $\mathbf{b}(N) = (b_1(N) \leq \cdots \leq b_N(N))$, we have that $\mathbf{a}(N) +_\theta \mathbf{b}(N)$ is the deterministic N-tuple $(b_1(N) + a \leq \cdots \leq b_N(N) + a)$, as mentioned right after Definition 1.2. Hence if $\lim_{N\to\infty} \mathbf{b}(N) \overset{m}{=} \{m_k\}_{k\geq 1}$, then $\lim_{N\to\infty} \mathbf{a}(N) +_\theta \mathbf{b}(N) \overset{m}{=} \{\widetilde{m}_k\}_{k\geq 1}$. Then the claim follows from Theorem 1.5.

Example 4.8. Let $\sigma^2 > 0$ be any positive number and consider the sequence of γ-cumulants:

$$\kappa_l^{\sigma^2} := \begin{cases} \sigma^2, & \text{if } l = 2, \\ 0, & \text{otherwise.} \end{cases} \tag{4.14}$$

Denote the corresponding sequence of moments as $\{m_k^{\sigma^2}\}_{k\geq 1} := T^\gamma_{\kappa\to m}(\{\kappa_l^{\sigma^2}\}_{l\geq 1})$. Observe that $\{m_k^{\sigma^2}\}_{k\geq 1}$ is the sequence of moments of a rescaled version of the distribution of Theorem 4.1, that we denote μ_{γ,σ^2}. From (4.14) and the definition of γ-convolution, it follows that for any $\sigma_1^2, \sigma_2^2 > 0$ we have $\mu_{\gamma,\,\sigma_1^2+\sigma_2^1} = \mu_{\gamma,\sigma_1^2} \boxplus_\gamma \mu_{\gamma,\sigma_2^2}$, or equivalently

$$\{m_k^{\sigma_1^2+\sigma_2^2}\}_{k\geq 1} = \{m_k^{\sigma_1^2}\}_{k\geq 1} \boxplus_\gamma \{m_k^{\sigma_2^2}\}_{k\geq 1}.$$

Example 4.9. Let $\lambda > 0$ be arbitrary and consider the constant sequence of γ-cumulants: $\kappa_l^\lambda := \lambda$, for all $l = 1, 2, \dots$. Denote the corresponding sequence of moments as $\{m_k^\lambda\}_{k\geq 1} := T^\gamma_{\kappa\to m}(\{\kappa_l^\lambda\}_{l\geq 1})$. It is known that $\{m_k^\lambda\}_{k\geq 1}$ is the sequence of moments of a probability measure ν_γ^λ. It follows that for any $\lambda_1, \lambda_2 > 0$, we have $\nu_\gamma^{\lambda_1+\lambda_2} = \nu_\gamma^{\lambda_1} \boxplus_\gamma \nu_\gamma^{\lambda_2}$, or equivalently

$$\{m_k^{\lambda_1+\lambda_2}\}_{k\geq 1} = \{m_k^{\lambda_1}\}_{k\geq 1} \boxplus_\gamma \{m_k^{\lambda_2}\}_{k\geq 1}.$$

The measure ν_γ^λ was studied in [ABMV13,TT19], where it was shown to be the limit of the empirical measures of beta Laguerre ensembles in the limit $N \to \infty$, $\beta N \to 2\gamma$, $N/M \to \lambda$. The density of ν_γ^λ can be obtained from [TT19, Lemma 2.1]; note that in that paper our parameters γ, λ are denoted by c, α, respectively. We also refer to [MP21, Section 5.4 and Figure 5] for additional details and plots of the densities. Since all the γ-cumulants of ν_γ^λ are equal to each other, this measure is similar to the Poisson and the Marchenko-Pastur distributions whose cumulants, respectively, free cumulants, are all the same.

4.4. Law of large numbers for ergodic measures. We start this section by providing some context in the complex case $\theta = 1$ (or $\beta = 2$). The infinite-dimensional unitary group $U(\infty)$ is defined as the union of the groups of $N \times N$ unitary matrices, $\bigcup_{N=1}^{\infty} U(N)$, where we embed $U(N)$ into $U(N+1)$ as the subgroup of operators fixing the $(N+1)$st basis vector. Each element of $U(\infty)$ is an infinite matrix, such that for some $N = 1, 2, \ldots$, its top $N \times N$ corner is unitary and outside this corner we have 1s on the diagonal and 0s everywhere else. Consider the space $\mathcal{H}$ of infinite complex Hermitian matrices with rows and columns parameterized by positive integers i and j. $U(\infty)$ acts on $\mathcal{H}$ by conjugations and one can ask about random matrices in $\mathcal{H}$ whose laws are invariant under such action. Their probability distributions form a simplex and much of the work on conjugation-invariant matrices comes down to study extreme points of this simplex—ergodic conjugation-invariant random matrices in $\mathcal{H}$. In [Pi91, OV96] these matrices were completely classified: they depend on a sequence of real parameters $\{\alpha_i\}_{i=1}^{\infty}$ with $\sum_{i=1}^{\infty} (\alpha_i)^2 \le \infty$ and two reals $\delta_1 \in \mathbb{R}$, $\delta_2 \ge 0$ and are given by an infinite sum:

$$\delta_1 \mathcal{I} + \sqrt{\delta_2}\, \frac{X + X^*}{2} + \sum_{i=1}^{\infty} \alpha_i \left(\tfrac{1}{2} V_i V_i^* - \mathcal{I} \right), \tag{4.15}$$

where $\mathcal{I}$ is the identical matrix (with 1s on the diagonal and 0s everywhere else), X is a matrix with i.i.d. Gaussian $\mathcal{N}(0,1) + \mathbf{i}\mathcal{N}(0,1)$ elements, V_i is an infinite (column-) vector with i.i.d. Gaussian $\mathcal{N}(0,1) + \mathbf{i}\mathcal{N}(0,1)$ components and all the involved matrices are independent. Note that if the only non-zero parameter is δ_2, then (4.15) gives the Gaussian Unitary Ensemble (a particular case of Wigner matrices). If the only non-zero parameters are $\alpha_1 = \alpha_2 = \cdots = \alpha_K = 1$ and $\delta_1 = K$, then (4.15) gives the Laguerre Unitary Ensemble (a particular case of Wishart or sample-covariance matrices).

As was first mentioned in [OV96, Remark 8.3] and recently studied in details in [AN19], the problem of classification of conjugation-invariant infinite complex Hermitian matrices has a general θ-version related to the θ-corners processes of Definition 2.1. Roughly speaking, while there are no infinite self-adjoint matrices in the general θ-version, one can make sense of the *distribution of eigenvalues* of the top-left principal submatrices (corners) of the infinite self-adjoint matrices. One of the problems addressed in [AN19] (see Theorem 1.13 there) is the classification of ergodic random matrices at general values of θ. Since there are no bona fide matrices, this problem actually asks for distributions of N-tuples, for $N = 1, 2, \ldots$, satisfying certain coherence relations—the distributions should be regarded as the eigenvalue distributions of the $N \times N$ corners of an ergodic matrix. It turns out that the set of parameters remains the same as in the $\theta = 1$ case. The law of the top-left 1×1 corner η of an ergodic matrix at general values of θ has characteristic function

$$\mathbb{E} e^{\mathbf{i}t\eta} = \mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(\mathbf{i}t), \quad \text{where} \quad \mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(z) := \exp\Bigl(\delta_1 z + \tfrac{\delta_2}{2\theta} z^2\Bigr) \cdot \prod_{i=1}^{\infty} \frac{\exp(-\alpha_i z)}{\bigl(1 - \frac{\alpha_i}{\theta} z\bigr)^{\theta}}. \tag{4.16}$$

More generally, there is a formula that uniquely determines the eigenvalue distribution of the corners, namely if $\eta_1 \le \eta_2 \le \cdots \le \eta_N$ are the random eigenvalues of the $N \times N$ corner, then their Bessel generating function is explicit:

$$\mathbb{E}\bigl[B_{(\eta_1,\ldots,\eta_N)}(x_1, \ldots, x_N;\, \theta)\bigr] = \prod_{j=1}^{N} \mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(x_j). \tag{4.17}$$

We will take (4.17) as our definition of the distributions on N-tuples $(\eta_1 \le \cdots \le \eta_N)$; these distributions are the ergodic measures of [OV96, AN19]. For them, we prove the following Law of Large Numbers in the regime $N \to \infty$, $\theta N \to \gamma$.

Theorem 4.10. *Suppose that θ and $\{\alpha_i\}_{i=1}^{\infty}$, δ_1, δ_2 vary with N in such a way that $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$ and*

$$\ln\left(\mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(z)\right) \longrightarrow F(z) = \sum_{l=1}^{\infty} \frac{\kappa_l}{l} z^l, \tag{4.18}$$

uniformly over a complex neighborhood of 0. Then the eigenvalues $(\eta_1, \dots, \eta_N)$ of the $N \times N$ corners of the corresponding general θ ergodic random matrix converge in the sense of moments (as in Definitions 1.4 or 3.1) to a probability distribution with γ-cumulants κ_l, i.e. its moments are found by the expression of Theorem 3.10.

Remark 4.11. Choosing $\delta_2 = \theta$, so that $\mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(z) = \exp(z^2/2)$, we recover the LLN for the Gaussian β-ensembles as $N \to \infty$, $\beta N \to 2\gamma$, as in Sect. 4.1. Choosing $\alpha_1 = \alpha_2 = \cdots = \alpha_M = \theta$, $\delta_1 = M\theta$ with $M = \lfloor \lambda N/\gamma \rfloor$, so that $\mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(z) = (1-z)^{-\theta\lfloor \lambda N/\gamma \rfloor} \to (1-z)^{-\lambda}$, we recover the LLN for the Laguerre β-ensembles as $N \to \infty$, $\beta N \to 2\gamma$ and $M/N \to \lambda/\gamma$. The limiting probability measure is ν_γ^λ, as described in Example 4.9.

Remark 4.12. The formula (4.16) has a multiplicative structure: a product of $\mathcal{F}_{\theta;\{\alpha_i\},\delta_1,\delta_2}(z)$ functions is again a function of the same type. This property leads to the limits in Theorem 4.10 being infinitely-divisible with respect to γ-convolution $\boxplus_\gamma$. This is in agreement with Examples 4.8 and 4.9, which show that the measures $\mu_\gamma^{\sigma^2}$ and ν_γ^λ are γ-infinitely-divisible.

Proof of Theorem 4.10. Combining (4.17) with (4.18), we conclude that the BGF of $(\eta_1, \dots, \eta_N)$ is γ-LLN appropriate in the sense of Definition 3.3. Hence, by Theorem 3.8, $(\eta_1, \dots, \eta_N)$ converge in the sense of moments and the asymptotic moments are recovered from the γ-cumulants κ_l by using the map $T^\gamma_{\kappa\to m}$. □

4.5. Limit of projections. We again start from the complex case $\theta = 1$. This time we fix $N = 1, 2, \dots$ and a deterministic N-tuple of reals $a_1 \le \cdots \le a_N$. Let A_N be a uniformly random $N \times N$ complex Hermitian matrix with eigenvalues $a_1, \dots, a_N$ and let A_m be the $m \times m$ top-left submatrix of A_N. We now fix $\tau > 1$, set $m = \lfloor N/\tau \rfloor$ and send $N \to \infty$. If we assume that the empirical measures of eigenvalues of A_N, $\frac{1}{N}\sum_{i=1}^N \delta_{a_i}$, converge to a limiting probability measure μ, then the (random) empirical measures of eigenvalues of A_m converge to a (deterministic) measure $\mu^{\boxplus\tau}$. For integer τ this measure is the same as the free convolution of τ copies of μ, hence, $\mu^{\boxplus\tau}$ can be called a fractional convolution power, see [STJ20] for a recent study and references.

We now present an analogue of the operation $\mu \mapsto \mu^{\boxplus\tau}$ in our $\theta \to 0$ asymptotic framework.

Theorem 4.13. *Fix real numbers $\gamma > 0$ and $\tau > 1$. Suppose that for each $N = 1, 2, \dots$, we are given an N-tuple of reals $a_1 \le \cdots \le a_N$, and let $\{y_i^k\}_{1\le i\le k\le N}$ be the θ-corners process with top row $a_1, \dots, a_N$, as in Definition 2.1. (In particular, this means*

$y_1^N = a_1, \ldots, y_N^N = a_N$.) *Define the empirical measures*

$$\rho_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{a_i}, \qquad \rho_N^\tau = \frac{1}{\lfloor N/\tau \rfloor} \sum_{i=1}^{\lfloor N/\tau \rfloor} \delta_{y_i^{\lfloor N/\tau \rfloor}},$$

and suppose that all measures ρ_N are supported inside a segment $[-C, C]$ and as $N \to \infty$, ρ_N weakly converge to a probability measure μ (supported inside the same segment). Then as $N \to \infty$, $\theta \to 0$ with $\theta N \to \gamma$, the (random) measures ρ_N^τ converge weakly, in probability to a deterministic measure $\mu^{\tau,\gamma}$. If $\{m_k\}_{k\geq 1}$ are the moments of μ and $\{m_k^\tau\}_{k\geq 1}$ are the moments of $\mu^{\tau,\gamma}$, then

$$T_{m\to\kappa}^{\gamma}\big(\{m_k\}_{k\geq 1}\big) = T_{m\to\kappa}^{\gamma/\tau}\big(\{m_k^\tau\}_{k\geq 1}\big). \tag{4.19}$$

In other words, γ-cumulants of μ coincide with $\frac{\gamma}{\tau}$-cumulants of $\mu^{\tau,\gamma}$.

Remark 4.14. The condition of support inside $[-C, C]$ is used to guarantee that all the involved measures are determined by their moments; it can be replaced by other uniqueness conditions for the moments problem.

Proof of Theorem 4.13. Convergence $\rho_N \to \mu$ and the condition on the support of ρ_N imply that the moments of ρ_N converge to those of μ. Hence, the sequence of delta-measures (unit masses) on N-tuples $(a_1, \ldots, a_N)$ satisfies LLN in the sense of Definition 3.1. Thus, Theorem 3.8 yields that it is γ-LLN appropriate, i.e., its BGF

$$G_{N;\theta}(x_1, \ldots, x_N) = B_{(a_1,\ldots,a_N)}(x_1, \ldots, x_N;\ \theta)$$

satisfies the conditions of Definition 3.3. Let $\tilde{G}_{N;\theta}$ denote the BGF of the $\lfloor N/\tau \rfloor$-tuple of reals $y_1^{\lfloor N/\tau \rfloor} \leq \cdots \leq y_{\lfloor N/\tau \rfloor}^{\lfloor N/\tau \rfloor}$. Then Definition 2.5 implies that

$$\tilde{G}_{N;\theta}(x_1, \ldots, x_{\lfloor N/\tau \rfloor}) = G_{N;\theta}(x_1, \ldots, x_{\lfloor N/\tau \rfloor}, 0, 0, \ldots, 0),$$

where there are $N - \lfloor N/\tau \rfloor$ in the right-hand side. Hence, the partial derivatives of $\ln(\tilde{G}_{N;\theta})$ coincide with partial derivatives of $\ln(G_{N;\theta})$ and, therefore, the former is γ-LLN appropriate. It is important to emphasize at this point that we use the same θ for $G_{N;\theta}$ and $\tilde{G}_{N;\theta}$, however, the number of variables for the latter is $\lfloor N/\tau \rfloor$ rather than N. This leads to γ being divided by τ. It remains to use Theorem 3.8 yet again to conclude that the random measures ρ_N^τ converge in the sense of moments and consequently also weakly, in probability. □

In general, we do not know any simple criteria on when a given sequence of numbers is a sequence of γ-cumulants corresponding to a probability measure. Yet Theorem 4.13 leads to an interesting comparison between different γ's.

Corollary 4.15. *Take a sequence of real numbers $\kappa_1, \kappa_2, \ldots$ and suppose that for some $\gamma_0 > 0$, these numbers are γ_0-cumulants of some probability measure ν, i.e., $\{\kappa_l\}_{l\geq 1} = T_{m\to\kappa}^{\gamma_0}\big(\{m_k\}_{k\geq 1}\big)$ with $m_k = \int_{\mathbb{R}} x^k \nu(dx)$. Then for each $0 < \gamma < \gamma_0$ the same numbers are also γ-cumulants of some probability measure. In particular, sending $\gamma \to 0$, we also have that the sequence $0!\kappa_1, 1!\kappa_2, 2!\kappa_3, \ldots$ gives conventional cumulants of some probability measure.*

In fact, $0!\kappa_1, 1!\kappa_2, 2!\kappa_3, \ldots$ are the conventional cumulants of the probability measure μ that is related to ν by means of the generalized Markov-Krein transform (3.10).

Proof. We apply Theorem 4.13 with $\tau = \gamma_0/\gamma$. The theorem was proven only for compactly supported measures, but we can approximate any measure by compactly supported ones. Finally, the convergence of γ-cumulants to conventional cumulants as $\gamma \to 0$ is discussed in Sect. 8. □

Remark 4.16. Theorem 4.13, or just equation (4.19), defines for $\tau > 1$ the (τ, γ)-*projection* map

$$\Pi_{\tau,\gamma} : \mu \mapsto \mu^{\tau,\gamma},$$

which maps the space of probability measures of compact support to itself. It would be interesting to study the possibility of an extension of (τ, γ)-projection map to all probability measures. In particular, the probability measures $\mu_\gamma^{\sigma^2}$ and ν_γ^λ from Examples 4.8 and 4.9 should map to the measures of the same type: $\Pi_{\tau,\gamma}(\mu_\gamma^{\sigma^2}) = \mu_{\gamma/\tau}^{\sigma^2}$ and $\Pi_{\tau,\gamma}(\nu_\gamma^\lambda) = \nu_{\gamma/\tau}^\lambda$.

5. Law of Large Numbers at High Temperature

In this section, we prove Theorem 3.8. Recall that the real parameter $\gamma > 0$ is fixed, and we are interested in the limit regime $N \to \infty$, $\theta \to 0$, and $\theta N \to \gamma$.

Let us recall some terminology about partitions of numbers (rather than set partitions of Sect. 3). A partition λ is a weakly decreasing sequence of nonnegative integers $\lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq 0)$, $\lambda_i \in \mathbb{Z}_{\geq 0}$, such that $\sum_{i=1}^\infty \lambda_i < \infty$. The latter sum is denoted $|\lambda|$ and is called the *size* of the partition λ. If λ is a partition of size k, we write $\lambda \vdash k$. The *length* $\ell(\lambda)$ of λ is defined as the number of strictly positive parts of λ.

The partitions are often identified with Young diagrams, in which λ_i become the row lengths. We also need column lengths $\lambda'_1 \geq \lambda'_2 \geq \dots$ defined by $\lambda'_j = |\{i \geq 1 \mid \lambda_i \geq j\}|$. In particular, $\lambda'_1 = \ell(\lambda)$.

5.1. The asymptotic expansion of Dunkl operators. If F is a smooth symmetric function of the N variables $x_1, \dots, x_N$, then its Taylor series expansion is also symmetric and we can write the k-th order approximation as

$$F(x_1, \dots, x_N) = \sum_{\lambda:\, |\lambda| \leq k,\ \ell(\lambda) \leq N} c_F^\lambda \cdot M_\lambda(\vec{x}) + O(\|x\|^{k+1}), \tag{5.1}$$

where the sum is over partitions λ of size at most k and length at most N. Finally, $M_\lambda(\vec{x})$ is the monomial symmetric function:

$$M_\lambda(\vec{x}) = \sum_{\substack{(d_1,\dots,d_N) \in \mathbb{Z}_{\geq 0}^N, \text{ such that} \\ \lambda \text{ is the rearrangement of } d_i \text{ in nonincreasing order}}} x_1^{d_1} \cdot x_2^{d_2} \cdots x_N^{d_N}.$$

Theorem 5.1. *Fix $k = 1, 2, \dots$ and a partition λ with $|\lambda| = k$. Let $F(x_1, \dots, x_N)$ be a symmetric function of $(x_1, \dots, x_N) \in \mathbb{R}^N$, which is $(k+1)$-times continuously*

differentiable in a neighborhood of $(0,\dots,0)$ *and satisfies* $F(0,\dots,0)=0$. *Then we have:*

$$N^{-\ell(\lambda)}\left[\prod_{i=1}^{\ell(\lambda)}\mathcal{P}_{\lambda_i}\right]\exp\big(F(x_1,\dots,x_N)\big)\Big|_{x_1=\cdots=x_N=0}=b_\lambda^\lambda\cdot c_F^\lambda+\sum_{\mu:\,|\mu|=k,\,\ell(\mu)>\ell(\lambda)}b_\mu^\lambda\cdot c_F^\mu$$
$$+L\Big(c_F^{(i)},\ 1\le i\le k-1\Big)+R_1\Big(c_F^\nu,\ |\nu|<k\Big)+N^{-1}R_2\Big(c_F^\nu,\ |\nu|\le k\Big),\tag{5.2}$$

where b_μ^λ *are coefficients, which are uniformly bounded in the regime* $N\to\infty$, $\theta\to 0$, $\theta N\to\gamma$. *In particular,*

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}}b_\lambda^\lambda=\prod_{i=1}^{\ell(\lambda)}\lambda_i(1+\gamma)_{\lambda_i-1}.\tag{5.3}$$

Further,

$$L\Big(c_F^{(i)},\ 1\le i\le k-1\Big)=\prod_{i=1}^{\ell(\lambda)}\Big([z^0](\partial+\gamma d+*_g)^{\lambda_i-1}g(z)\Big)-k(1+\gamma)_{k-1}c_F^{(k)}\mathbf{1}_{\ell(\lambda)=1},\tag{5.4}$$

where (m) *is the one-row Young diagram of size* m, *the operators* ∂, d, $*_g$ *are the ones introduced in Definition 3.6, and* $g(z):=\sum_{n=1}^\infty nc_F^{(n)}z^{n-1}$.

Next, $R_1\Big(c_F^\nu,\ |\nu|<k\Big)$ *is a polynomial in* $c_F^\nu,\ |\nu|<k$, *such that:*

- *If we assign the degree* $|\nu|$ *to each* c_F^ν, *then* R_1 *is homogeneous of degree* k.
- *The coefficients of the monomials in* R_1 *are uniformly bounded in the regime* $N\to\infty$, $\theta\to 0$, $\theta N\to\gamma$.
- *Each monomial in* R_1 *has at least one factor* c_F^ν *with* $\ell(\nu)>1$.

Finally, $R_2\Big(c_F^\nu,\ |\nu|\le k\Big)$ *is a homogeneous polynomial in* $c_F^\nu,\ |\nu|\le k$, *of degree* k *and with uniformly bounded coefficients (in the same regime).*

Before proving Theorem 5.1, let us use it to deliver the proof of Theorem 3.8.

Proof of Theorem 3.8. First, take a LLN-appropriate sequence $\{\mu_N\}_N$ with associated sequence of real numbers $\{\kappa_l\}_{l\ge1}$. Let $\{m_k\}_{k\ge1}$ be the image of $\{\kappa_l\}_{l\ge1}$ under the map $T^\gamma_{\kappa\to m}$, that is, each m_k is the function of the κ_l's given by (3.1). We aim to show that $\{\mu_N\}_N$ satisfies a LLN with associated sequence of real numbers $\{m_k\}_{k\ge1}$.

Let us denote the BGF of μ_N by $G_{N;\theta}(x_1,\dots,x_N)$. Let $s=1,2,\dots$ and $k_1,\dots,k_s\in\mathbb{Z}_{\ge1}$ be arbitrary. By Proposition 2.11 (or Proposition 2.13 for distributions), we have

$$\mathbb{E}_{\mu_N}\left[\prod_{i=1}^s p_{k_i}^N\right]=N^{-s}\left(\prod_{i=1}^s\mathcal{P}_{k_i}\right)G_{N;\theta}(x_1,\dots,x_N)\Big|_{x_1=\cdots=x_N=0}.\tag{5.5}$$

Without loss of generality, we assume that $k_1\ge k_2\ge\cdots\ge k_s$, so that the k_i's form a partition. Since $G_{N;\theta}$ is holomorphic in a neighborhood of the origin and $G_{N;\theta}(0,\dots,0)=1$, then there is a holomorphic function $F_{N;\theta}(x_1,\dots,x_N)$ in a neighborhood of the origin such that $G_{N;\theta}=\exp(F_{N;\theta})$ and $F_{N;\theta}(0,\dots,0)=0$. The

functions $F_{N;\theta}(x_1, \dots, x_N)$ are smooth and symmetric in the real variables $x_1, \dots, x_N$, so we can consider their Taylor expansions:

$$F_{N;\theta}(x_1, \dots, x_N) = \sum_{\lambda:\, |\lambda| \le k,\, \ell(\lambda) \le N} c^{\lambda}_{F_{N;\theta}} \cdot M_\lambda(\vec{x}) + O(\|x\|^{k+1}).$$

By LLN-appropriateness,

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} c^{(n)}_{F_{N;\theta}} = \frac{\kappa_n}{n}, \qquad \lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} c^{\mu}_{F_{N;\theta}} = 0, \quad \text{if } \ell(\mu) > 1.$$

Apply Theorem 5.1 to the function $F_{N;\theta}$ and the partition $\lambda = (k_1 \ge \cdots \ge k_s)$. Let us take the limit of each term in the resulting right-hand side of (5.2) in the limit regime $N \to \infty, \theta \to 0, \theta N \to \gamma$:

- In the first line, if $s > 1$, then each term involves some $c^{\mu}_{F_{N;\theta}}$ with $\ell(\mu) > 1$, and therefore tends to 0. Otherwise, if $s = 1$, then there is a single asymptotically non-vanishing term, namely $b^{(k_1)}_{(k_1)} \cdot c^{(k_1)}_{F_{N;\theta}}$, which converges to $(1+\gamma)_{k_1-1}\,\kappa_{k_1}$.
- The polynomial R_1 converges to 0, since each of its monomials involves some $c^{\mu}_{F_{N;\theta}}$ with $\ell(\mu) > 1$ and, therefore, vanishes asymptotically.
- The polynomial $\frac{1}{N} R_2$ converges to 0 due to the $\frac{1}{N}$ prefactor.
- The polynomial L converges to

$$\prod_{i=1}^{s} \Big([z^0](\partial + \gamma d + *_g)^{k_i - 1} g(z) \Big) - (1+\gamma)_{k_1-1} \kappa_{k_1} \mathbf{1}_{s=1},$$

where $g(z) = \sum_{n=1}^{\infty} \kappa_n z^{n-1}$, due to the fact that the power series $\sum_{n=1}^{\infty} n c^{(n)}_{F_{N;\theta}} z^{n-1}$ converges coefficient-wise to $g(z)$.

Combining the terms coming from the above four items, we conclude that

$$\lim_{N\to\infty} \mathbb{E}_{\mu_N} \left[\prod_{i=1}^{s} p^N_{k_i} \right] = \prod_{i=1}^{s} \Big([z^0](\partial + \gamma d + *_g)^{k_i - 1} g(z) \Big).$$

We have thus arrived at the Law of Large Numbers with m_k given by (3.1).

In the opposite direction, take a sequence μ_N which satisfies the Law of Large Numbers with associated sequence $\{m_k\}_{k\ge 1}$. Let $\{\kappa_l\}_{l\ge 1}$ be the image of $\{m_k\}_{k\ge 1}$ under the map $T^{\gamma}_{m\to\kappa}$. Again let $G_{N;\theta}(x_1, \dots, x_N)$ be the BGF of μ_N. We show that μ_N is LLN-appropriate with corresponding sequence $\{\kappa_l\}_{l\ge 1}$, that is, we are going to establish the conditions on partial derivatives of Definition 3.3. This will be done by induction on the total order of the derivative. For the inductive step, we assume that for all $s \le k-1$, the asymptotic behavior of all partial derivatives of order s is already established, i.e. we assume that the limits

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_s}} \ln\left(G_{N;\theta}\right)\Big|_{x_1=\cdots=x_N=0}, \quad i_1, \dots, i_s \in \mathbb{Z}_{\ge 1},$$

exist and are equal to zero unless $i_1 = \cdots = i_s$, in which case the limit is equal to $(s-1)! \cdot \kappa_s$.

Our task is to prove the two conditions of Definition 3.3 for $\ell = k$ and for $r = k$. Let $p(k)$ be the total number of partitions of k and consider the $p(k)$ expressions (5.2) obtained by making λ run over all partitions of k and letting $F_N := F_{N;\theta}$ be determined through $G_{N;\theta} = \exp(F_{N;\theta})$. We regard the left-hand sides and the coefficients $c^{\mu}_{F_{N;\theta}}$ with $|\mu| < k$, as constants, while we regard the terms $c^{\lambda}_{F_{N;\theta}}$ with $|\lambda| = k$, as variables; then we can treat these expressions as $p(k)$ linear equations for the $p(k)$ variables $c^{\lambda}_{F_{N;\theta}}$ with $|\lambda| = k$. The coefficients of these equations generally depend on N and θ, and moreover we know the $N \to \infty, \theta \to \infty, \theta N \to \infty$ asymptotic behavior of the left-hand sides of (5.2) as well as L and R_1 in the right-hand side (by the inductive hypothesis). The form of the first line of (5.2) implies that the matrix of coefficients of these equations becomes triangular as $N \to \infty, \theta \to 0, \theta N \to \gamma$ in the lexicographic order $\le$ on partitions of size k, viewed as vectors of column lengths $(\lambda'_1, \lambda'_2, \dots)$, because $\ell(\mu) > \ell(\lambda)$ implies $\mu > \lambda$. The diagonal elements have nonzero limits, because of (5.3).

We can rewrite these linear equations in the matrix notation. Let $B^{N,\theta}$ be the $p(k) \times p(k)$ matrix with matrix elements

$$B^{N,\theta}(\mu, \lambda) = b^{\lambda}_{\mu}.$$

Further, let $\mathbf{c}^N$ denote the $p(k)$-dimensional column-vector with coordinates $c^{\lambda}_{F_{N;\theta}}$, $|\lambda| = k$. Then the previous paragraph can be summarized as a matrix equation

$$B^{N,\theta} \cdot \mathbf{c}^N = \mathbf{r}^N, \tag{5.6}$$

where the vector $\mathbf{c}^N$ is unknown and the right-hand side $\mathbf{r}^N$ is known. The key property of (5.6) is that the entries of the inverse matrix $(B^{N,\theta})^{-1}$ are bounded as $N \to \infty$, $\theta \to \infty$, $\theta N \to \gamma$; this follows from triangularity of $B^{N,\theta}$ and non-zero limits for its diagonal entries. Let $\mathbf{c}^{\infty}$ denote another $p(k)$-dimensional vector, in which the first coordinate (corresponding to the one-row partition (k)) is $\frac{\kappa_k}{k}$ (here κ_k is found from (3.2), in which the numbers $m_1, m_2, \dots$ are known us) and all other coordinates are zeros. The first part of the proof (where we showed that each LLN-appropriate sequence satisfies LLN) and the induction hypothesis imply that

$$B^{N,\theta} \cdot \mathbf{c}^{\infty} = \mathbf{r}^N + o(1), \tag{5.7}$$

where $o(1)$ is a vanishing term as $N \to \infty, \theta \to \infty, \theta N \to \gamma$. Multiplying (5.6) and (5.7) by $(B^{N,\theta})^{-1}$ and comparing the results, we conclude that

$$\lim_{\substack{N\to\infty,\theta\to 0,\\ \theta N\to\gamma}} \mathbf{c}^N = \mathbf{c}^{\infty}.$$

□

5.2. Proof of Theorem 5.1. We start by reducing to the case of F being a symmetric polynomial.

Lemma 5.2. *Suppose that F is a $(k+1)$-times continuously differentiable function in a neighborhood of $(0, \dots, 0) \in \mathbb{C}^N$, with Taylor expansion (5.1). Then for any λ with $|\lambda| = k$, we have*

$$\left[\prod_{i=1}^{\ell(\lambda)} \mathcal{P}_{\lambda_i}\right] \exp\bigl(F(x_1, \dots, x_N)\bigr)\Big|_{x_1=\dots=x_N=0} = \left[\prod_{i=1}^{\ell(\lambda)} \mathcal{P}_{\lambda_i}\right] \exp\bigl(\tilde{F}(x_1, \dots, x_N)\bigr)\Big|_{x_1=\dots=x_N=0},$$

where

$$\tilde{F}(x_1, \dots, x_N) = \sum_{\nu:\, |\nu| \le k} c_F^{\nu} \cdot M_{\nu}(\vec{x}).$$

Proof. We have

$$\exp\big(F(x_1, \dots, x_N)\big) = \exp\big(\tilde{F}(x_1, \dots, x_N)\big) + R(x_1, \dots, x_N),$$

where R is a $(k+1)$-times continuously differentiable function, satisfying $R = O(\|x\|^{k+1})$ as $(x_1, \dots, x_N) \to (0, \dots, 0)$. It remains to show that after we apply k operators of the form $\frac{\partial}{\partial x_i}$ or $\frac{1-s_{ij}}{x_i - x_j}$ to R, the resulting function $R^{(k)}$ is continuous and vanishes at $(0, \dots, 0)$. For that we let $R^{(m)}$, $m = 1, 2, \dots, k$ be the result of application of m such operators and prove by induction in m that $R^{(m)}$ is $(k+1-m)$-times continuously differentiable and satisfies $R^{(m)} = O(\|x\|^{k+1-m})$. The induction step is proven by applying to $R^{(m)}$ the Taylor's theorem with remainder in the integral form. □

By virtue of Lemma 5.2, we can (and will) assume for the remainder of this section that

$$F(x_1, \dots, x_N) = \sum_{\nu:\, |\nu| \le k} c_F^{\nu} \cdot M_{\lambda}(\vec{x}).$$

Next, consider any product of k operators, each of which is either $\frac{\partial}{\partial x_i}$ for some i, or $\frac{1}{x_i - x_j}(1 - s_{ij})$ for some i and j. We apply these operators inductively to $\exp(F)$, using the following rules:

$$\begin{aligned} &\frac{\partial}{\partial x_i}\big[H(x_1, \dots, x_N) \cdot \exp(F(x_1, \dots, x_N))\big] \\ &\quad = \left(\frac{\partial}{\partial x_i} H(x_1, \dots, x_N) + H(x_1, \dots, x_N) \frac{\partial}{\partial x_i} F(x_1, \dots, x_N)\right) \cdot \exp(F(x_1, \dots, x_N)), \end{aligned} \tag{5.8}$$

$$\begin{aligned} &\frac{1}{x_i - x_j}(1 - s_{ij})\big[H(x_1, \dots, x_N) \cdot \exp(F(x_1, \dots, x_N))\big] \\ &\quad = \left(\frac{1}{x_i - x_j}(1 - s_{ij}) H(x_1, \dots, x_N)\right) \cdot \exp(F(x_1, \dots, x_N)). \end{aligned} \tag{5.9}$$

Hence, taking into account that $F(0, \dots, 0) = 0$, the result of acting by such product on $\exp(F)$ and then setting all variables equal to 0 is a finite linear combination of products of actions of $\frac{\partial}{\partial x_i}$ and $\frac{1}{x_i - x_j}(1 - s_{ij})$ on the function F, and then picking up the constant term of the polynomial. Since F is a polynomial with coefficients c_F^{λ} and the actions of $\frac{\partial}{\partial x_i}$ and $\frac{1}{x_i - x_j}(1 - s_{ij})$ on monomials are clear, we conclude the following statement.

Lemma 5.3. *For any k indices $1 \le i_1, \dots, i_k \le N$, the expression*

$$\left(\prod_{m=1}^{k} \mathcal{D}_{i_m}\right) \exp(F(x_1 \dots, x_N))\Big|_{x_1 = \dots = x_N = 0} \tag{5.10}$$

is a homogeneous polynomial of degree k in c_F^{λ} (if we regard each c_F^{λ} as a degree $|\lambda|$ variable), whose coefficients are uniformly bounded as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$.

Proof. By definition, each $\mathcal{D}_i$ is linear combination of N terms, each of which is $\frac{\partial}{\partial x_i}$ or $\frac{1}{x_i - x_j}(1 - s_{ij})$. Observe that any of these two simple operators decreases the degree of a polynomial in the variables $x_1, \dots, x_N$ by 1. Therefore, using the rules (5.8) and (5.9), the expression (5.10) is a polynomial in the coefficients of the degree k component of $\exp(\sum_{\lambda:\,|\lambda|\le k} c_F^\lambda \cdot M_\lambda(\vec{x}))$. Such polynomial is therefore in the variables c_F^λ and is homogeneous of degree k, because of how we assigned the degrees to the c_λ^F's.

In the formula (2.1) for $\mathcal{D}_i$, the term $\frac{\partial}{\partial x_i}$ comes with unit coefficient, and the remaining terms $\frac{1}{x_i - x_j}(1 - s_{ij})$ come with a prefactor θ, which decays as γ/N as $N \to \infty$. Hence, expanding $\prod_{m=1}^k \mathcal{D}_{i_m}$ as a linear combination of products of the operators $\frac{\partial}{\partial x_i}$ and $\frac{1}{x_i - x_j}(1 - s_{ij})$, we see that the coefficients of the polynomial (5.10) in the variables c_λ^F are uniformly bounded in the regime of our interest. □

Corollary 5.4. *Take any partition λ with $|\lambda| = k$. As $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$, we have*

$$\begin{aligned} N^{-\ell(\lambda)} \left[\prod_{i=1}^{\ell(\lambda)} \mathcal{P}_{\lambda_i} \right] \exp\bigl(F(x_1, \dots, x_N)\bigr)\Big|_{x_1=\dots=x_N=0} \\ = \left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i} \right] \exp\bigl(F(x_1, \dots, x_N)\bigr)\Big|_{x_1=\dots=x_N=0} \\ + N^{-1} R_3, \end{aligned} \tag{5.11}$$

where R_3 is a homogeneous polynomial of degree k in the coefficients c_F^ν (if we regard each c_F^ν as a degree $|\nu|$ variable), and with uniformly bounded coefficients.

Proof. Each $\mathcal{P}_{\lambda_i}$ is a sum of N terms $(\mathcal{D}_j)^{\lambda_i}$, $j = 1, \dots, N$. Hence, $\left[\prod_{i=1}^{\ell(\lambda)} \mathcal{P}_{\lambda_i}\right]$ is a sum of $N^{\ell(\lambda)}$ terms, each of which is a finite (independent of N and θ) product of $(\mathcal{D}_j)^{\lambda_i}$. For all but $O(N^{\ell(\lambda)-1})$ of these terms, the indices j are all distinct. Hence, by symmetry of F, the result of the action of such product on $\exp(F)$ is the same as that of $\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}$, after setting all variables x_i equal to zero. Dividing by $N^{-\ell(\lambda)}$, we get the desired statement. □

For the rest of the section, we analyze $\left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \exp\bigl(F(x_1, \dots, x_N)\bigr)\Big|_{x_1=\dots=x_N=0}$. In view of Corollary 5.4 we need to show that it has an expansion of the form of the right-hand side of (5.2).

Proposition 5.5. *Take any partition λ with $|\lambda| = k$. We have*

$$\begin{aligned} \left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i} \right] \exp\bigl(F(x_1, \dots, x_N)\bigr)\Big|_{x_1=\dots=x_N=0} = b_\lambda^\lambda \cdot c_F^\lambda \\ + \sum_{\mu:\,|\mu|=k,\ \ell(\mu)>\ell(\lambda)} b_\mu^\lambda \cdot c_F^\mu + R + O\left(\frac{1}{N}\right), \end{aligned}$$

where the coefficients b^λ_μ are uniformly bounded in the regime $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$. In particular,

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} b^\lambda_\lambda = \prod_{i=1}^{\ell(\lambda)} \lambda_i (1+\gamma)_{\lambda_i - 1}.$$

Moreover, R is a homogeneous polynomial of degree k in the coefficients c^ν_F with $|\nu| < k$, i.e., it does not involve the coefficients c^ν_F with $|\nu| = k$. Finally, $O\left(\frac{1}{N}\right)$ stands for a linear polynomial in the coefficients c^ν_F with $|\nu| = k$, whose coefficients are of the order $O\left(\frac{1}{N}\right)$, as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$.

Proof. By Lemma 5.3, $\left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \exp\bigl(F(x_1,\dots,x_N)\bigr)\Big|_{x_1=\dots=x_N=0}$ is a homogeneous polynomial of degree k in the coefficients c^ν_F, $|\nu| \le k$. Hence, its linear component is of the form

$$\sum_{\mu:\,|\mu|=k} b^\lambda_\mu \cdot c^\mu_F.$$

Therefore, two steps remain:

1. We need to show that $b^\lambda_\mu = O\left(\frac{1}{N}\right)$ unless $\ell(\mu) > \ell(\lambda)$ or $\mu = \lambda$.
2. We need to find the limit of b^λ_λ as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$.

We first claim that the part of $\left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \exp\bigl(F(x_1,\dots,x_N)\bigr)\Big|_{x_1=\dots=x_N=0}$ involving the coefficients c^μ_F with $|\mu| = k$ is given by

$$\left[\prod_{i=2}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \cdot \left[\mathcal{D}_1^{\lambda_1 - 1}\right] \frac{\partial}{\partial x_1} F(x_1,\dots,x_N)\Big|_{x_1=\dots=x_N=0}. \tag{5.12}$$

Indeed, the operators $\mathcal{D}_i$ commute, hence, we can apply $\mathcal{D}_1$ first. In the very first application of $\mathcal{D}_1$, the terms $\frac{1}{x_1 - x_j}(1 - s_{1j})$ can be omitted, since $(1 - s_{1j})$ annihilates the symmetric function $\exp(F)$. Hence, the result of the first application of $\mathcal{D}_1$ is $\frac{\partial F}{\partial x_1} \cdot \exp(F)$. Using formula (5.8), we see that all the next applications of partial derivatives $\frac{\partial}{\partial x_1}$ should never act on $\exp(F)$, as otherwise we are not getting the terms c^μ_F with $|\mu| = k$. Similarly, when we further apply $\prod_{i=2}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}$, we should not act on $\exp(F)$. Hence, we can omit $\exp(F)$, as it does not contribute to the computation. Therefore we get (5.12).

We analyze (5.12) by using the expansion $F(x_1,\dots,x_N) = \sum_{\mu:\,|\mu|\le k} c^\mu_F \cdot M_\mu(\vec{x})$ in monomials and looking at each monomial separately. Note that each operator $\mathcal{D}_i$ lowers by 1 the degree of the monomial on which it acts. Since we apply $\frac{\partial}{\partial x_1}$, then $k-1$ operators $\mathcal{D}_i$, and then plug in all variables equal to 0, the only way to get a non-zero contribution is by acting on a monomial of degree k. We conclude that the coefficient b^λ_μ is computed by

$$b^\lambda_\mu = \left[\prod_{i=2}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \cdot \left[\mathcal{D}_1^{\lambda_1 - 1}\right] \frac{\partial}{\partial x_1} M_\mu(\vec{x})\Big|_{x_1=\dots=x_N=0} = \left[\prod_{i=2}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \cdot \left[\mathcal{D}_1^{\lambda_1 - 1}\right] \frac{\partial}{\partial x_1} M_\mu(\vec{x}). \tag{5.13}$$

Each $\mathcal{D}_i$ is a sum of N operators. Hence, the operator in (5.13) can be represented as a sum of N^{k-1} operators, each of which is a product of the factors $\frac{\partial}{\partial x_i}$ and $\frac{\theta}{x_i - x_j}(1 - s_{ij})$.

Claim A. Only the terms in which all indices j are distinct and are all larger than $\ell(\lambda)$ contribute to the leading term of (5.13). All others combine together into a remainder of order $O\left(\frac{1}{N}\right)$.

For example, if $\lambda = (2, 1)$ (so that $\ell(\lambda) = 2$) and $\mu = (1, 1, 1)$, then (5.13) contains terms of the following types:

$$\text{(I)}\ \frac{\partial}{\partial x_2}\frac{\partial^2}{\partial x_1^2} M_{(1,1,1)}(\vec{x}), \qquad \text{(II)}\ \frac{\theta}{x_2 - x_j}(1 - s_{2j})\frac{\partial^2}{\partial x_1^2} M_{(1,1,1)}(\vec{x}), \quad j \neq 2,$$

$$\text{(III)}\ \frac{\partial}{\partial x_2}\frac{\theta}{x_1 - x_k}(1 - s_{1k})\frac{\partial}{\partial x_1} M_{(1,1,1)}(\vec{x}), \quad k \neq 1,$$

$$\text{(IV)}\ \frac{\theta}{x_2 - x_j}(1 - s_{2j})\frac{\theta}{x_1 - x_k}(1 - s_{1k})\frac{\partial}{\partial x_1} M_{(1,1,1)}(\vec{x}), \quad j \neq 2,\ k \neq 1.$$

Then Claim A states that the term (II) with $j = 1$, the term (III) with $k = 2$, and the terms (IV) with $j = k$ or $j = 1$ or $k = 2$, all combined give a contribution which is smaller, by a factor of N, than the contribution of all other terms of these four types, i.e. those with $j, k > 2$ and $j \neq k$.

Claim A is proven by a simple counting argument. Indeed, the terms with distinct indices are the generic ones: the number of terms where two indices coincide is smaller, by a factor of N, than the number of similar terms with distinct indices.

Next, using Claim A, let us take a look at the first application of $\mathcal{D}_2$ after we computed $\mathcal{D}_1^{\lambda-1}\frac{\partial}{\partial x_1} M_\mu(\vec{x})$. We could either apply $\frac{\partial}{\partial x_2}$ or we can apply $\frac{\theta}{x_2 - x_j}(1 - s_{2j})$. But due to symmetry in x_2 and x_j the result of the application of the latter operator vanishes. Hence, we have to use $\frac{\partial}{\partial x_2}$. Similarly, in the first application of $\mathcal{D}_3$ we need to use $\frac{\partial}{\partial x_3}$, etc. We conclude that

$$b_\mu^\lambda = \left[\mathcal{D}_{\ell(\lambda)}^{\lambda_{\ell(\lambda)}-1}\frac{\partial}{\partial x_{\ell(\lambda)}}\right]\cdots\left[\mathcal{D}_2^{\lambda_2-1}\frac{\partial}{\partial x_2}\right]\cdot\left[\mathcal{D}_1^{\lambda_1-1}\frac{\partial}{\partial x_1}\right] M_\mu(\vec{x}) + O\left(\frac{1}{N}\right). \tag{5.14}$$

We analyze the last expression in three steps.

Step 1. Let us show that if $\ell(\mu) < \ell(\lambda)$, then (5.14) is $O\left(\frac{1}{N}\right)$. Indeed, if $\ell(\mu) < \ell(\lambda)$, then each monomial in $M_\mu(\vec{x})$ is missing one of the variables $x_1, \ldots, x_{\ell(\lambda)}$. Say, it does not have x_m. Then, using the above Claim A, we see that when we apply $\frac{\partial}{\partial x_m}$ in (5.14), the expression has no dependence on x_m and, hence, the derivative vanishes.

Step 2. If $\ell(\mu) > \ell(\lambda)$, then b_μ^λ are bounded as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$, by Lemma 5.3 and we do not need to prove anything else about them.

Step 3. It remains to study the case $\ell(\mu) = \ell(\lambda) = \ell$. Let us expand $M_\mu(\vec{x})$ in monomials. If a monomial is missing one of the variables $x_1, \ldots, x_\ell$, then by the argument of Step 1, it does not contribute to b_μ^λ. Hence, since $\ell(\mu) = \ell(\lambda)$, it remains to study the monomials which involve $x_1, x_2, \ldots, x_\ell$ and no other variables.

Note that for $1 \le i \le \ell < j$, we have, using the degree-lowering operators (2.2)

$$\frac{\theta}{x_i - x_j}(1 - s_{ij})[x_1^{n_1} \cdots x_\ell^{n_\ell}] = \theta \left[\prod_{a \neq i} x_a^{n_a}\right] \frac{x_i^{n_i} - x_j^{n_i}}{x_i - x_j}$$
$$= \theta \left[\prod_{a \neq i} x_a^{n_a}\right] \left(x_i^{n_i - 1} + x_i^{n_i - 2} x_j + \cdots + x_j^{n_i - 1}\right) = \theta d_i [x_1^{n_1} \cdots x_\ell^{n_\ell}] + x_j \cdot P, \tag{5.15}$$

where P is a polynomial of degree $n_1 + \cdots + n_\ell - 2$.

Using the above Claim A, one sees that if a factor x_j, $j > \ell$ appears in a monomial, then this factor cannot be annihilated by applying any operator $\frac{\partial}{\partial x_i}$, $i \leq \ell$, or any operator $\frac{\theta}{x_i - x_{j'}}(1 - s_{ij'})$, $i \leq \ell$, $j \neq j'$, unless this application makes the entire monomial vanish. Hence, the only way to get a non-zero contribution is by using the d_i term, but not the $x_j \cdot P$ term in (5.15). Thus, up to $O\left(\frac{1}{N}\right)$ error, the desired b_μ^λ can be alternatively computed as:

$$b_\mu^\lambda = \left[(\partial_\ell + \theta(N-1)d_\ell)^{\lambda_\ell - 1}\, \partial_\ell\right] \cdots \left[(\partial_1 + \theta(N-1)d_1)^{\lambda_1 - 1} \partial_1\right] M_\mu(x_1, \ldots, x_\ell) + O\left(\frac{1}{N}\right).$$

(Above we denoted $\frac{\partial}{\partial x_i}$ by ∂_i for all i.) The last operator lowers the degree of x_1 by λ_1, lowers the degree of x_2 by λ_2,..., lowers the degree of x_ℓ by λ_ℓ. Since $\lambda_1 + \cdots + \lambda_\ell = \mu_1 + \cdots + \mu_\ell$, the only way to get a non-zero contribution after these lowerings is by having $\lambda = \mu$. Therefore, $b_\mu^\lambda = O(N^{-1})$ if $\ell(\mu) = \ell(\lambda)$ and $\mu \neq \lambda$.

Finally, in the case $\mu = \lambda$, we have

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} b_\lambda^\lambda = \lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \left[(\partial_\ell + (N-1)\theta d_\ell)^{\lambda_\ell - 1}\, \partial_\ell\right] \cdots$$
$$\left[(\partial_1 + (N-1)\theta d_1)^{\lambda_1 - 1} \partial_1\right] x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_\ell^{\lambda_\ell}$$
$$= \left[(\partial_\ell + \gamma d_l)^{\lambda_\ell - 1}\, \partial_\ell\right] \cdots \left[(\partial_1 + \gamma d_1)^{\lambda_1 - 1} \partial_1\right] x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_\ell^{\lambda_\ell}$$
$$= \prod_{i=1}^{\ell} \lambda_i(\lambda_i - 1 + \gamma)(\lambda_i - 2 + \gamma) \cdots (1 + \gamma).$$

□

Proposition 5.5 gives the linear part, that is, the first line in (5.2). The next step is to identify $L(\cdot)$ in the second line of (5.2).

Proposition 5.6. *Take any partition λ with $|\lambda| = k$. We have*

$$\left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \exp\big(F(x_1, \ldots, x_N)\big)\Big|_{x_1 = \cdots = x_N = 0} = \prod_{i=1}^{\ell(\lambda)} \left([z^0](\partial + \gamma d + *_g)^{\lambda_i - 1} g(z)\right) + R + O\left(\frac{1}{N}\right), \tag{5.16}$$

where $g(z) = \sum_{n=1}^{\infty} n c_F^{(n)} z^{n-1}$, and ∂, d, $*_g$ are the operators from Definition 3.7. Moreover, R is a homogeneous polynomial of degree k in c_F^ν with $|\nu| \leq k$, such that each

monomial in it involves at least one ν with $\ell(\nu) > 1$. Finally, $O\left(\frac{1}{N}\right)$ is a (homogeneous of degree k) polynomial in c_F^{ν} with $|\nu| \le k$, whose coefficients are $O\left(\frac{1}{N}\right)$ as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$.

Proof. We only need to figure out the monomials which involve $c_F^{(n)}$, $n = 1, \dots, k$, and no other coefficients, so we are only interested in the following part of the left-hand side of (5.16) with

$$\left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \exp\left(\sum_{n=1}^{k} c_F^{(n)} M_{(n)}(\vec{x})\right)\Big|_{x_1=\dots=x_N=0}$$
$$= \left[\prod_{i=1}^{\ell(\lambda)} (\mathcal{D}_i)^{\lambda_i}\right] \prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n\right)\Big|_{x_1=\dots=x_N=0}. \tag{5.17}$$

Next, we recall that each $\mathcal{D}_i$ is a sum of N operators, so that the operator in (5.17) is a sum of N^k operators, each of which is a product of the factors $\frac{\partial}{\partial x_i}$ and $\frac{\theta}{x_i - x_j}(1 - s_{ij})$. As in Claim A in the proof of Proposition 5.5, we can and will assume without loss of generality that all indices j are distinct and larger than $\ell(\lambda)$—we only accumulate $O\left(\frac{1}{N}\right)$ error by making such assumption.

There are two consequences of this. First, like in (5.14), at this point for each i the very first application of $\mathcal{D}_i$ can be replaced by $\frac{\partial}{\partial x_i}$, since the operators $\frac{\theta}{x_i - x_j}(1 - s_{ij})$ act by 0 due to symmetry in i and j. Second, the operators no longer interact with each other in any way and the expression factorizes. This reasoning is very similar to that in the proof of Proposition 5.5, so we do not dwell on the details. As a result, up to $O\left(\frac{1}{N}\right)$ error, (5.17) is equal to

$$\prod_{i=1}^{\ell(\lambda)} \left((\mathcal{D}_i)^{\lambda_i - 1} \frac{\partial}{\partial x_i} \left[\prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n\right)\right]\Big|_{x_1=\dots=x_N=0} \right). \tag{5.18}$$

It remains to study the factor in (5.18) corresponding to a single i; without loss of generality, let us consider the case $i = 1$. We would like to understand

$$(\mathcal{D}_1)^{l-1} \frac{\partial}{\partial x_1} \left[\prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n\right)\right]\Big|_{x_1=\dots=x_N=0}$$
$$= (\mathcal{D}_1)^{l-1} \left[g(x_1) \prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n\right)\right]\Big|_{x_1=\dots=x_N=0}, \tag{5.19}$$

where

$$g(x_1) := \sum_{n=1}^{k} n c_F^{(n)} (x_1)^{n-1}.$$

Note that for any polynomial H we have

$$\frac{\partial}{\partial x_1}\left[H \cdot \prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n\right)\right] = \left(\frac{\partial H}{\partial x_1} + H \cdot g(x_1)\right) \cdot \prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n\right) \tag{5.20}$$

and

$$\frac{\theta}{x_1 - x_j}(1 - s_{1j})\left[H \cdot \prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n \right) \right]$$
$$= \left[\frac{\theta}{x_1 - x_j}(1 - s_{1j}) H \right] \cdot \prod_{t=1}^{N} \exp\left(\sum_{n=1}^{k} c_F^{(n)} (x_t)^n \right). \tag{5.21}$$

Combining (5.20) and (5.21), we can rewrite (5.19) as

$$(\mathcal{D}_1 + *_g)^{l-1} g(x_1)\Big|_{x_1 = \cdots = x_N = 0}, \tag{5.22}$$

where $*_g$ is the operator if multiplication by $g(x_1)$. It remains to note we can replace each operator $\frac{\theta}{x_1 - x_j}(1 - s_{1j})$ in $\mathcal{D}_1$ by θd_1. Indeed, this is done by the exact same reasoning that we used in the proof of Proposition 5.5, see (5.15). After we make this replacement, we conclude that (up to another $O\left(\frac{1}{N}\right)$ error) (5.22) and (5.19) are equal to

$$\left(\frac{\partial}{\partial x_1} + \gamma d_1 + *_g \right)^{l-1} g(x_1)\Big|_{x_1 = \cdots = x_N = 0} = [z^0](\partial + \gamma d + *_g)^{l-1} g(z).$$

Plugging this expression back into (5.18) gives the desired result. □

After all these preparations, it remains to put everything together, as follows.

Proof of Theorem 5.1. By Lemma 5.3 and Corollary 5.4, the left-hand side of (5.2) is a homogeneous polynomial in c_F^ν, $\nu \le k$, of degree k (if we regard each c_F^ν as a variable of degree $|\nu|$) with uniformly bounded coefficients as $N \to \infty$, $\theta \to 0$, $\theta N \to \gamma$. Proposition 5.5 identifies the linear part of this polynomial (corresponding to c_F^ν with $|\nu| = k$) with the first line in the right-hand side of (5.2). Proposition 5.6 identifies the polynomial L from the second line of the right-hand side of (5.2) and from (5.4). It remains to note that subtraction of $k(1+\gamma)_{k-1} c_F^{(k)}$ in (5.4) corresponds to the situation when the parts of the polynomial given by Propositions 5.5 and 5.6 overlap. □

6. From γ-Cumulants to Moments

The goal of this section is to prove Theorem 3.10. That is, let us begin with any real sequence $\kappa_1, \kappa_2, \ldots$, consider the power series $g(z) := \sum_{l=1}^{\infty} \kappa_l z^{l-1}$, and the operators ∂, d, and $*_g$ from Definition 3.6. We denote the constant term of a power series $h(z)$ by $[z^0]h(z)$.

Recall that $\mathscr{P}(k)$ denotes the collection of all set partitions of $[k]$, and for each $\pi \in \mathscr{P}(k)$ we introduced the γ-weight $W(\pi)$ in Definition 3.9. If $\pi = B_1 \sqcup \cdots \sqcup B_m \in \mathscr{P}(k)$, then the B_i's are called the *blocks* of π. The cardinality of the block B_i is denoted by $|B_i|$.

With these recollections, Theorem 3.10 says that for any $k \in \mathbb{Z}_{\ge 1}$, we must have the equality

$$[z^0](\partial + *_g + \gamma d)^{k-1}(g(z)) \stackrel{?}{=} \sum_{\pi = B_1 \sqcup \cdots \sqcup B_m \in \mathscr{P}(k)} W(\pi) \prod_{i=1}^{m} \kappa_{|B_i|}. \tag{6.1}$$

This relation and actually a more general version (see Theorem 6.2) will be proved in this section.

6.1. A refined combinatorial theorem. Let $a_1, a_2, \ldots$ and $\kappa_1, \kappa_2, \ldots$ be two arbitrary sequences of real numbers.

Definition 6.1. For any $k \in \mathbb{Z}_{\geq 1}$ and $\pi \in \mathscr{P}(k)$, we define the quantity $w(\pi)$, that will be called the *refined γ-weight of π* as follows.[11] Suppose that π has m blocks and label them $B_1, \ldots, B_m$ in such a way that the smallest element from B_i is smaller than the smallest element from B_j (hence, also smaller than all other elements from B_j), whenever $i < j$. Then define

$$w(\pi) := W(\pi) \cdot \kappa_{|B_1|} \prod_{i=2}^{m} a_{|B_i|}.$$

From the formula of $W(\pi)$ in Definition 3.9, we can give an expanded formula for the refined γ-weight $w(\pi)$ of the set partition $\pi = B_1 \sqcup \cdots \sqcup B_m$. Recall that the values $p(i), q(i)$, $1 \leq i \leq m$, are defined by

$$\begin{aligned} p(i) &:= \#\{j \in \{1, \ldots, |B_i| - 1\} \mid \{b_j^i + 1, \ldots, b_{j+1}^i - 1\} \cap B_t \neq \emptyset, \\ &\qquad \text{for some block } B_t \text{ with } t < i\}, \\ q(i) &:= |B_i| - 1 - p(i). \end{aligned}$$

In particular, $p(1) = 0, q(1) = |B_1| - 1$. The quantity $p(i)$ can be computed by the graphical procedure described in Sect. 3.2.

Define the *weight $w(B_i)$ of the block B_i with respect to π* by

$$w(B_i) := \begin{cases} (\gamma + 1)_{|B_1|-1} \cdot \kappa_{|B_1|}, & \text{if } i = 1, \\ p(i)! \cdot (\gamma + p(i) + 1)_{q(i)} \cdot a_{|B_i|}, & \text{if } i \geq 2. \end{cases} \tag{6.2}$$

For example, if $i \geq 2$ and B_i is a singleton, then $p(i) = q(i) = 0$ and $w(B_i) = a_1$.

The refined γ-weight of the set partition $\pi = B_1 \sqcup \cdots \sqcup B_m$ then equals

$$\begin{aligned} w(\pi) &= \prod_{i=1}^{m} w(B_i) \\ &= \big((\gamma + 1)_{|B_1|-1} \cdot \kappa_{|B_1|}\big) \cdot \prod_{i=2}^{m} \big(p(i)! \cdot (\gamma + p(i) + 1)_{q(i)} \cdot a_{|B_i|}\big). \end{aligned} \tag{6.3}$$

Observe that under the identifications $a_i \mapsto \kappa_i$, for all i, we have

$$w(\pi)|_{a_i \mapsto \kappa_i} = W(\pi) \prod_{i=1}^{m} \kappa_{|B_i|}.$$

This is why we call $w(\pi)$ the *refined* γ-weight of π.

We show the refined γ-weights for the same examples of set partitions given in Sect. 3.2. The set partition $\{1, 2, 5, 7\} \sqcup \{3, 4, 6\} \in \mathscr{P}(7)$ graphically shown in Fig. 1 has refined γ-weight

$$w(\pi) = (\gamma + 1)(\gamma + 2)^2(\gamma + 3) \cdot \kappa_4 a_3.$$

[11] We omit the dependence on γ and on the sequences $a_1, a_2, \ldots$ and $\kappa_1, \kappa_2, \ldots$ from the notation $w(\pi)$.

For the set partition $\{1,4\} \sqcup \{2,6\} \sqcup \{3,5,7\} \in \mathscr{P}(7)$ shown in Fig. 2, the refined γ-weight is

$$w(\pi) = 2(\gamma+1) \cdot \kappa_2 a_2 a_3.$$

As a final example, the set partition $\{1,3,4,5,6\} \sqcup \{2,7\} \in \mathscr{P}(7)$ shown in Fig. 3 has refined γ-weight

$$w(\pi) = (\gamma+1)(\gamma+2)(\gamma+3)(\gamma+4) \cdot \kappa_5 a_2.$$

Theorem 6.2. *Set*

$$g(z) = \sum_{l=1}^{\infty} \kappa_l z^{l-1}, \qquad a(z) = \sum_{l=1}^{\infty} a_l z^{l-1}.$$

Then we have

$$[z^0](\partial + *_a + \gamma d)^{k-1}(g(z)) = \sum_{\pi = B_1 \sqcup \cdots \sqcup B_m \in \mathscr{P}(k)} w(\pi). \tag{6.4}$$

On the left side, we have the constant term of a power series. On the right side, the sum ranges over set partitions of $[k]$*, and the refined* γ*-weight* $w(\pi)$ *is the one introduced in Definition 6.1.*

Note that this result implies Theorem 3.10: indeed, we apply Theorem 6.2 and set $a_i = \kappa_i$. In the rest of this section we prove Theorem 6.2.

6.2. Preliminary lemmas.

Lemma 6.3. *Let* $x, y \in \mathbb{Z}_{\geq 0}$ *be arbitrary, and let* z *be any complex number. Then*

$$(y+1) \sum_{i=1}^{x} (z+i)_y = (z+x)_{y+1} - (z)_{y+1}. \tag{6.5}$$

Proof. The proof is induction on x. If $x = 0$, then both sides of (6.5) vanish. The difference of the left-hand sides of (6.5) at $x = t$ and $x = t-1$ is $(y+1)(z+t)_y$. The difference of the right-hand sides of (6.5) is the same:

$$(z+t)_{y+1} - (z+t-1)_{y+1} = (z+t)_y \big(z+t+y - (z+t-1)\big) = (z+t)_y \cdot (y+1).$$

□

For a sequence of 0s and 1s, a *descent* is defined as a substring 10 in this sequence. Let $\mathrm{des}(\zeta)$ denote the number of descents in a 0-1 sequence ζ. For instance, $\mathrm{des}(1100) = 1$ and $\mathrm{des}(0101010) = 3$.

Lemma 6.4. *For any two integers* $N \geq 1$ *and* $0 \leq M \leq N$ *and any* $\gamma \in \mathbb{R}$*, we have*

$$\sum_{\substack{\zeta = (\zeta_1, \ldots, \zeta_N) \in \{0,1\}^N \\ \sum_{i=1}^N \zeta_i = M}} \mathrm{des}(\zeta)! \cdot (\gamma + \mathrm{des}(\zeta) + 1)_{M - \mathrm{des}(\zeta)} = (\gamma + 1 + N - M)_M. \tag{6.6}$$

Proof. Let $K(N, M, d)$ denote the total number of sequences $\zeta \in \{0, 1\}^N$ with $\sum_{i=1}^N \zeta_i = M$ and $\mathrm{des}(\zeta) = d$. With this notation, we would like to prove that:

$$\sum_{d=0}^{M} K(N, M, d) \cdot d! \cdot (\gamma + d + 1)_{M-d} = (\gamma + 1 + N - M)_M, \quad N \geq 1, \quad 0 \leq M \leq N. \tag{6.7}$$

Our proof is induction on N. If $N = 1$, then both sides of (6.7) are 1 at $M = 0$ and both sides are $(\gamma + 1)$ at $M = 1$. For the induction step, assume that (6.7) holds for all values $\leq N$, and let us prove it for $N + 1$, and an arbitrary $0 \leq M \leq N + 1$. We notice that the statement is straightforward at $M = 0$. If $M > 0$, then we use the following recurrence for $K(N, M, d)$, which is obtained by considering the position of the right-most 1 in a sequence ζ:

$$K(N + 1, M, d) = K(N, M - 1, d) + \sum_{p=1}^{N-M+1} K(N - p, M - 1, d - 1). \tag{6.8}$$

If $d = M$, then the first term in (6.8) in not needed; If $d = 0$, then the second term in (6.8) is not needed. Hence, the left-hand side of (6.7) for N replaced with $N + 1$ can be rewritten using (6.8) as

$$\begin{aligned} &\sum_{d=0}^{M} K(N + 1, M, d) \cdot d! \cdot (\gamma + d + 1)_{M-d} \\ &= \sum_{d=0}^{M-1} K(N, M - 1, d) \cdot d! \cdot (\gamma + d + 1)_{M-d} \\ &+ \sum_{d=1}^{M} \sum_{p=1}^{N-M+1} K(N - p, M - 1, d - 1) \cdot d! \cdot (\gamma + d + 1)_{M-d}. \end{aligned} \tag{6.9}$$

For the first sum in the right-hand side of (6.9), we use $d! \cdot (\gamma + d + 1)_{M-d} = d! \cdot (\gamma + d + 1)_{M-1-d} \cdot (\gamma + M)$ and induction assumption to evaluate it as

$$(\gamma + 2 + N - M)_{M-1} \cdot (\gamma + M). \tag{6.10}$$

For the second sum in the right-hand side of (6.9) we change the order of summation and evaluate the sum over d for fixed p using the induction assumption and identity (valid for $d > 0$)

$$d! \cdot (\gamma + d + 1)_{M-d} = (d - 1)! \cdot (\gamma + d)_{M-d} \cdot (\gamma + M) - (d - 1)! \cdot (\gamma + 1 + d)_{M-d} \cdot \gamma.$$

Hence, the pth term evaluates to

$$(\gamma + 2 + N - p - M)_{M-1} \cdot (\gamma + M) - (\gamma + 3 + N - p - M)_{M-1} \cdot \gamma. \tag{6.11}$$

Combining (6.10) with (6.11), we transform (6.9) to

$$(\gamma+M)\sum_{p=0}^{N-M+1}(\gamma+2+N-p-M)_{M-1}-\gamma\sum_{p=1}^{N-M+1}(\gamma+3+N-p-M)_{M-1}$$
$$=M\sum_{p=0}^{N-M}(\gamma+2+N-p-M)_{M-1}+(\gamma+M)\cdot(\gamma+1)_{M-1}.$$

We compute the last sum using Lemma 6.3, resulting in

$$(\gamma+2+N-M)_M-(\gamma+1)_M+(\gamma+M)\cdot(\gamma+1)_{M-1}=(\gamma+2+N-M)_M.$$

□

6.3. Proof of Theorem 6.2. The proof is by induction on k.

Step 1. For the base cases, let us consider $k=1$ and $k=2$.

Case $k=1$. Clearly, $[z^0]g(z)=[z^0](\kappa_1+\kappa_2 z+\dots)=\kappa_1$. On the other hand, there is exactly one set partition of $[1]$, namely $\pi=\{1\}$ with $B_1=\{1\}$. For this partition, $p(1)=q(1)=0$, so it follows that $w(\pi)=\kappa_1$, as needed.

Case $k=2$. Here, $[z^0](\partial+*_a+\gamma d)g(z)=[z^0](g'(z)+g(z)a(z)+\gamma dg(z))=\kappa_2+\kappa_1a_1+\gamma\kappa_2=\kappa_1a_1+(\gamma+1)\kappa_2$. On the other hand, there are two set partitions of $[2]$, namely $\pi_1=\{1\}\sqcup\{2\}$ with $B_1=\{1\}$, $B_2=\{2\}$, and $\pi_2=\{1,2\}$ with $B_1=\{1,2\}$. For π_1, $p(1)=p(2)=q(1)=q(2)=0$, so $w(\pi_1)=\kappa_1a_1$. For π_2, $p(1)=0$ and $q(1)=1$, so $w(\pi_2)=(\gamma+1)\kappa_2$. Therefore, $w(\pi_1)+w(\pi_2)=\kappa_1a_1+(\gamma+1)\kappa_2$.

Step 2. Suppose that the statement of the theorem is true for certain $k\geq 2$. For the induction step, we prove the statement for $k+1$, i.e. we aim to obtain the formula for $[z^0](\partial+*_a+\gamma d)^k(g(z))$.

Let $b_1,b_2,b_3,\dots$ be the quantities defined by

$$(\partial+*_a+\gamma d)(g(z))=b_1+b_2z+b_3z^2+\cdots.$$

From the definition of the operators ∂, $*_a$, and d, we have

$$b_n=\sum_{j=1}^{n}\kappa_{n+1-j}a_j+(\gamma+n)\kappa_{n+1},\quad n\in\mathbb{Z}_{\geq 1}. \tag{6.12}$$

Next, use the induction hypothesis to obtain the combinatorial formula:

$$\begin{aligned}[z^0](\partial+*_a+\gamma d)^k(g(z))&=[z^0](\partial+*_a+\gamma d)^{k-1}(b_1+b_2z+b_3z^2+\dots)\\&=\sum_{\tilde\pi=\tilde B_1\sqcup\cdots\sqcup\tilde B_m\in\mathscr{P}(k)}\tilde w(\tilde\pi),\end{aligned} \tag{6.13}$$

where $\tilde w(\tilde\pi)$ is given in the theorem (see (6.3)), but instead of the κ_i's, we should use the b_i's:

$$\tilde w(\tilde\pi)=\Big((\gamma+1)_{|\tilde B_1|-1}\cdot b_{|\tilde B_1|}\Big)\cdot\prod_{i=2}^{m}\Big(p(i)!\cdot(\gamma+p(i)+1)_{q(i)}\cdot a_{|\tilde B_i|}\Big). \tag{6.14}$$

From (6.12), this equals

$$\tilde{w}(\tilde{\pi}) = \sum_{j=1}^{|\tilde{B}_1|} \Big((\gamma+1)_{|\tilde{B}_1|-1} \cdot \kappa_{|\tilde{B}_1|+1-j}\, a_j \Big) \cdot \prod_{i=2}^{m} \Big(p(i)! \cdot (\gamma + p(i) + 1)_{q(i)} \cdot a_{|\tilde{B}_i|} \Big) + \Big((\gamma+1)_{|\tilde{B}_1|-1} \cdot (\gamma + |\tilde{B}_1|) \kappa_{|\tilde{B}_1|+1} \Big) \cdot \prod_{i=2}^{m} \Big(p(i)! \cdot (\gamma + p(i) + 1)_{q(i)} \cdot a_{|\tilde{B}_i|} \Big). \tag{6.15}$$

Our next goal is to obtain a different combinatorial expression for (6.13), (6.15)—we should get the right-hand side of (6.4) for $k+1$, namely a formula that involves set partitions of $[k+1]$.

Step 3. Given a set partition π of $[k+1] = \{1, 2, 3, \ldots, k+1\}$, consider the set partition $\tilde{\pi}$ of $\{1, 3, 4, \ldots, k+1\}$ that is obtained from π by taking the union of the blocks that contain 1 and 2, and then removing 2. If $\tilde{\pi}$ is obtained from π in this fashion, we say that π *maps to* $\tilde{\pi}$ and denote this relation by $\pi \to \tilde{\pi}$. Observe that if 1 and 2 belong to the same block of π, then π and $\tilde{\pi}$ have the same number of blocks. On the other hand, if 1 and 2 belong to different blocks of π, then $\tilde{\pi}$ has one block fewer than π. For instance, for $k = 2$ we have 5 set partitions of $\{1, 2, 3\}$ which are mapped to two set partitions of $\{1, 3\}$:

$$\{1\} \sqcup \{2, 3\} \to \{1, 3\}, \qquad \{1, 3\} \sqcup \{2\} \to \{1, 3\}, \qquad \{1, 2, 3\} \to \{1, 3\},$$
$$\{1\} \sqcup \{2\} \sqcup \{3\} \to \{1\} \sqcup \{3\}, \qquad \{1, 2\} \sqcup \{3\} \to \{1\} \sqcup \{3\}.$$

For a set partition $\tilde{\pi}$ of $\{1, 3, 4, \ldots, k+1\}$ we define the numbers $p(i)$, $q(i)$ and the weight $\tilde{w}(\tilde{\pi})$ by identifying $\{1, 3, 4, \ldots, k+1\}$ with $\{1, 2, \ldots, k\}$ in a monotone way and using the previous formula (6.14). Note that essentially nothing changes in the definition, as the way we compute the numbers $p(i)$, $q(i)$ and the weight $\tilde{w}(\cdot)$ depends only on the order of the elements of the set that we are partitioning rather than the labels of these elements. Hence, we use the same $\tilde{w}(\tilde{\pi})$ notation no matter whether $\tilde{\pi}$ is a partition of $\{1, 3, 4, \ldots, k+1\}$ or $\tilde{\pi}$ is a partition of $\{1, 2, \ldots, k\}$.

Our goal now is to prove that for each set partition $\tilde{\pi}$ of $\{1, 3, 4, \ldots, k+1\}$ we have an identity:

$$\sum_{\substack{\pi \in \mathscr{P}(k+1) \\ \pi \to \tilde{\pi}}} w(\pi) \stackrel{?}{=} \tilde{w}(\tilde{\pi}). \tag{6.16}$$

The last equation together with (6.13) implies the induction step. We fix $\tilde{\pi}$ and let its blocks be $\tilde{B}_1, \ldots, \tilde{B}_m$ ordered, as before, by their minimal elements, so that $\tilde{B}_1$ contains 1. We calculate the sum (6.16) by splitting the terms into several subsets. Define $T \subseteq \mathscr{P}(k+1)$ as the subset of those set partitions π, mapped to $\tilde{\pi}$, for which 1 and 2 belong to the same block of π. Next, for any $r \in \{1, 2, \ldots, |\tilde{B}_1|\}$, define $T_r \subseteq \mathscr{P}(k+1)$ as the subset of those set partitions π, mapped to $\tilde{\pi}$, for which 1 and 2 belong to distinct blocks of π and the block where 2 belongs is of size r. The sets T and $T_1, \ldots, T_{|\tilde{B}_1|}$ are all disjoint; they depend on $\tilde{\pi}$, but we omit this dependence from the notations. With these notations the desired identity (6.16) is rewritten

$$\sum_{\pi \in T \cup T_1 \cup T_2 \cup \cdots \cup T_{|\tilde{B}_1|}} w(\pi) \stackrel{?}{=} \tilde{w}(\tilde{\pi}). \tag{6.17}$$

In step 4 below, we prove that $\sum_{\pi\in T} w(\pi)$ is equal to the second line of (6.15). In steps 5 and 6, we prove that for any $r \in \{1, 2, \dots, |\tilde{B}_1|\}$ the sum $\sum_{\pi\in T_r} w(\pi)$ is equal to the $j = r$ summand in (6.15). After these steps are done, the identity (6.17) would follow and thus the proof would be complete.

Step 4. In this step we calculate $\sum_{\pi\in T} w(\pi)$. Recall that T contains all set partitions π of $[k + 1]$ that map to $\tilde{\pi}$ and such that 1 and 2 belong to the same block of π. In fact, for a given $\tilde{\pi}$ with blocks $\tilde{B}_1, \dots \tilde{B}_m$, there is only one set partition in T: it is $\pi = B_1 \sqcup \dots \sqcup B_m$, where $B_1 = \tilde{B}_1 \cup \{2\}$ and $B_h = \tilde{B}_h$, $h = 2, 3, \dots, m$. Since 1 and 2 are adjacent in the ordering of $[k + 1]$ and they belong to the same block of π, it is clear that the weight of B_i, $i \geq 2$, in the computation of $w(\pi)$ is the same as the weight of $\tilde{B}_i$ in the computation of $\tilde{w}(\tilde{\pi})$. The weight of B_1 in the computation of $w(\pi)$ is $(\gamma + 1)_{|B_1|-1} \cdot \kappa_{|B_1|} = (\gamma + 1)_{|\tilde{B}_1|} \cdot \kappa_{|\tilde{B}_1|+1}$. As a result,

$$w(\pi) = \Big((\gamma + 1)_{|\tilde{B}_1|} \cdot \kappa_{|\tilde{B}_1|+1}\Big) \cdot \prod_{i=2}^{m} \Big(p(i)! \cdot (\gamma + p(i) + 1)_{q(i)} \cdot a_{|\tilde{B}_i|}\Big).$$

This is exactly the second line of (6.15), as desired.

Step 5. In this step we calculate $\sum_{\pi\in T_r} w(\pi)$, $1 \leq r \leq |\tilde{B}_1|$, in the case when $\tilde{\pi}$ is a one-block set partition, $\tilde{\pi} = \tilde{B}_1 = \{1, 3, 4, \dots, k + 1\}$ and $|\tilde{B}_1| = k$. Hence, set partitions π in T_r have two blocks $\pi = B_1 \sqcup B_2$, $1 \in B_1, 2 \in B_2$ and $|B_2| = r$. Therefore, $|B_1| = k + 1 - r$.

We can identify elements of T_r with 0–1 sequences $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_{k-1})$ of length $k - 1$ and with $\sum_{i=1}^{k-1} \zeta_i = k - r$ through:

$$B_1(\zeta) = \{1\} \cup \{i + 2 \mid \zeta_i = 1\}, \quad B_2(\zeta) = \{2\} \cup \{i + 2 \mid \zeta_i = 0\}.$$

In words, the positions where $\zeta_i = 1$ encode the elements of B_1 from the set $\{3, 4, \dots, k+1\}$. With this notation, we rewrite

$$\sum_{\pi\in T_r} w(\pi) = \sum_{\substack{\zeta\in\{0,1\}^{k-1} \\ \sum_{i=1}^{k-1}\zeta_i = k-r}} w\big(B_1(\zeta) \sqcup B_2(\zeta)\big).$$

Let us compute $w\big(B_1(\zeta) \sqcup B_2(\zeta)\big)$. By definition (6.2), the weight of block B_1 is

$$w(B_1) = (\gamma + 1)_{k-r}\, \kappa_{k-r+1}.$$

Further, note that in the notations of Lemma 6.4, $p(2) = \mathrm{des}(\zeta)$: Indeed, in the arc diagrams (as in Figs. 1, 2, 3) a roof of the block B_2 with no legs intersecting it corresponds to a substring 00 in ζ, whereas a roof with $m \geq 1$ legs intersecting it corresponds to a substring $011\cdots110$ (with m ones), and this substring gives exactly one descent in ζ. Hence, by definition (6.2), we have

$$w(B_2) = \mathrm{des}(\zeta)!(\gamma + \mathrm{des}(\zeta) + 1)_{r-\mathrm{des}(\zeta)-1} a_r.$$

We obtain

$$\sum_{\pi\in T_r} w(\pi) = \sum_{\substack{\zeta\in\{0,1\}^{k-1} \\ \sum_{i=1}^{k-1}\zeta_i = k-r}} (\gamma + 1)_{k-r}\, \kappa_{k-r+1} \cdot \mathrm{des}(\zeta)!\, (\gamma + \mathrm{des}(\zeta) + 1)_{r-\mathrm{des}(\zeta)-1}\, a_r$$

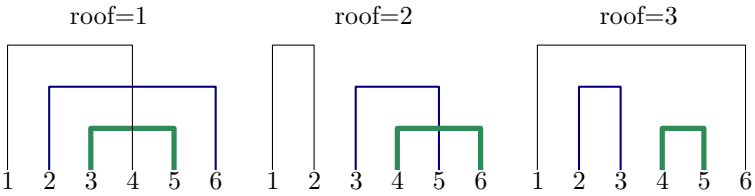

Fig. 4. Perfect matchings of [6] with three possible values of roof(π)

$$= (\gamma+1)_{k-r}\,\kappa_{k-r+1}\,a_r \sum_{\substack{\zeta\in\{0,1\}^{k-1}\\ \sum_{i=1}^{k-1}\zeta_i=k-r}} \mathrm{des}(\zeta)!\,(\gamma+\mathrm{des}(\zeta)+1)_{k-r-\mathrm{des}(\zeta)}\cdot\frac{(\gamma+1)_{r-1}}{(\gamma+1)_{k-r}}$$

$$= (\gamma+1)_{r-1}\,\kappa_{k-r+1}a_r\,(\gamma+r)_{k-r} = (\gamma+1)_{k-1}\kappa_{k-r+1}a_r, \tag{6.18}$$

where we used Lemma 6.4 with $N = k-1$ and $M = k-r$ for the equality between the second and the third lines of (6.18). Since $|\tilde{B}_1| = k$ and there are no $\tilde{B}_h$ with $h > 1$, the third line in (6.18) matches the $j = r$ term in (6.15), as desired (Fig. 4).

Step 6. We now extend the computation of Step 5 and calculate $\sum_{\pi\in T_r} w(\pi)$, $1 \le r \le |\tilde{B}_1|$, for arbitrary $\tilde{\pi} = \tilde{B}_1 \sqcup \cdots \sqcup \tilde{B}_m$. By definition of T_r each set partition $\pi \in T_r$ has $m+1$ blocks $B_1, \ldots, B_{m+1}$ and we have $1 \in B_1$, $2 \in B_2$, $|B_2| = r$, and $B_i = \tilde{B}_{i-1}$, $3 \le i \le m+1$. The key observation for this step is that for $i \ge 3$ the weight of the block B_i, $w(B_i)$, is the same as the weight of the block $\tilde{B}_{i-1}$, $\tilde{w}(\tilde{B}_{i-1})$: this is because the blocks B_i and $\tilde{B}_i$ coincide as sets and the legs intersecting their roofs in the arc diagrams also coincide. Hence, we have

$$w(\pi) = w(B_1)w(B_2)\prod_{i=2}^{m}\Big(p(i)!\cdot(\gamma+p(i)+1)_{q(i)}\cdot a_{|\tilde{B}_i|}\Big)$$

It remains to sum the last formula over all possible choices of B_1 and B_2. Since $B_1 \cup B_2 = \tilde{B}_1 \cup \{2\}$ is fixed, this is the same computation as in Step 5, but with k replaced by $|\tilde{B}_1|$. As a result, we get

$$\sum_{\pi\in T_r} w(\pi) = (\gamma+1)_{|\tilde{B}_1|-1}\kappa_{|\tilde{B}_1|-r+1}a_r\prod_{i=2}^{m}\Big(p(i)!\cdot(\gamma+p(i)+1)_{q(i)}\cdot a_{|\tilde{B}_i|}\Big),$$

which matches the $j = r$ term in (6.15), as desired.

7. From Moments to γ-Cumulants

Let $\{m_k\}_{k\ge1}$ and $\{\kappa_l\}_{l\ge1}$ be real sequences related by $\{\kappa_l\}_{l\ge1} = T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\ge1})$. In this section, we prove Theorem 3.11, namely the following identity:

$$\exp\left(\sum_{l=1}^{\infty}\frac{\kappa_l y^l}{l}\right) \stackrel{?}{=} [z^0]\left(\sum_{n=0}^{\infty}\frac{(yz)^n}{(\gamma)_n}\right)\exp\left(\gamma\sum_{k=1}^{\infty}\frac{m_k}{kz^k}\right). \tag{7.1}$$

The central idea of our proof is to apply Theorem 3.8 to the measure μ_N which is the Dirac delta-mass at a single N-tuple $(a_1^{(N)}, \ldots, a_N^{(N)})$. The Bessel generating function of

μ_N is the multivariate Bessel function $B_{(a_1^{(N)},\dots,a_N^{(N)})}(x_1,\dots,x_N;\theta)$, which allows us to use the known formulas for $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$ and get the asymptotic expressions for the partial derivatives of the logarithm of the BGF at 0. We remark that it would be interesting to find a more direct combinatorial proof, explaining how (7.1) matches the expressions of Theorem 3.10.

As our proof of Theorem 3.11 is based on the asymptotic analysis of $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$, we start by collecting formulas for this function. Assume, as usual, that $a_1 \le a_2 \le \cdots \le a_N$ are real and $y \in \mathbb{C}$. There are at least three different ways to think about $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$:

1. The Taylor series expansion for $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$ (which is a limit of the binomial formula for Jack polynomials of [OO97]) reads

$$B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)=\sum_{k=0}^{\infty}\frac{Q_{(k)}(a_1,\dots,a_N;\,\theta)}{(\theta N)_k}y^k, \tag{7.2}$$

   where $Q_{(k)}(a_1,\dots,a_N;\,\theta)$ is the value of the N-variable Jack symmetric polynomial (with normalization as for Q-functions in [M98, Chapter VI, Section 10] or [OO97]) parameterized by one-row partition (k) at the point $(a_1,\dots,a_N)$. The expansion (7.2) is a particular case of [OO97, (4.2)]; that article uses the same parameter θ for the Jack polynomials, but it is worth mentioning that some other authors (e.g., [S89] or [M98, Section VI.10]) use $\alpha=\theta^{-1}$ instead.
2. The contour integral representation for $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$ claims that for any complex y with $\Re y>0$ we have

$$B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)=\frac{\Gamma(\theta N)}{y^{\theta N-1}}\frac{1}{2\pi\mathbf{i}}\int_{\mathscr{C}_\infty}\exp(yz)\prod_{j=1}^{N}(z-a_j)^{-\theta}dz, \tag{7.3}$$

   where the infinite contour $\mathscr{C}_\infty$ in this formula is positively oriented and is formed by the segment $[M-r\mathbf{i},\,M+r\mathbf{i}]$ and the horizontal lines $[M+r\mathbf{i},\,-\infty+r\mathbf{i})$, $[M-r\mathbf{i},\,-\infty-r\mathbf{i})$, for real numbers $M>a_N$ and $r>0$. The proof of (7.3) can be found in [Cu19, Theorem 5.1] and the same article contains a complementary integral representation for $\Re y<0$.
3. A stochastic representation for $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$ reads

$$B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)=\mathbb{E}\left[\exp\left(y\sum_{i=1}^{N}a_i\eta_i\right)\right], \tag{7.4}$$

   where $(\eta_1,\dots,\eta_N)$ is a Dirichlet-distributed random vector with all parameters equal to θ. The proof of (7.4) can be found in [AN19, Proposition 5.1], although in some forms this statement was known before, see, e.g., [OV96, Remark 8.3].

Either of the above three approaches can be used to establish a formula for the generating function of derivatives of $B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)$ at 0:

Proposition 7.1. *For any $\theta>0$, $N\in\mathbb{Z}_{>0}$, $y\in\mathbb{C}$, and $a_1\le a_2\le\cdots\le a_N$ we have the expansion*

$$B_{(a_1,\dots,a_N)}(y,0^{N-1};\theta)=\sum_{k=0}^{\infty}\frac{c_k}{(\theta N)_k}y^k, \tag{7.5}$$

where the numbers c_k are found from the following Taylor series expansion:

$$\sum_{k=0}^{\infty} c_k z^k = \prod_{i=1}^{N} (1 - a_i z)^{-\theta}. \tag{7.6}$$

The series (7.5) is uniformly convergent over y in compact subsets of $\mathbb{C}$.

Proof. According to (7.2), the coefficients c_k in (7.5) are computed as $c_k = Q_{(k)}(a_1, \ldots, a_N; \theta)$. The generating function for the one-row Jack polynomials is well-known, see [M98, Section VI.10, top formula on page 378] or [S89, (9)]. We have:

$$\sum_{k=0}^{\infty} Q_{(k)}(a_1, \ldots, a_N; \theta) z^k = \prod_{i=1}^{N} (1 - a_i z)^{-\theta},$$

which proves (7.6). Finally, uniform convergence of (7.5) follows either from the fact that we deal with a Taylor series expansion of an entire function, or from bounds $c_k < Cr^k$ for some $C > 0, r > 0$, which can be extracted from (7.6). □

In view of Theorem 3.8, the desired identity (7.1) of Theorem 3.11 becomes the limit as $N \to \infty, \theta N \to \gamma$ of Theorem 7.1, as we now explain.

Proof of Theorem 3.11. **Step 1.** We first show that the formulas (3.7) and (3.8) are equivalent. Indeed, (3.7) says that the coefficient of y^n in $\exp\left(\sum_{l=1}^{\infty} \frac{\kappa_l}{l} y^l\right)$ can be computed as the constant term of

$$\frac{z^n}{(\gamma)_n} \exp\left(\gamma \sum_{k=1}^{\infty} \frac{m_k}{k} z^{-k}\right).$$

On the other hand, (3.8) says that the same coefficient can be computed as $\frac{1}{(\gamma)_n}$ times the coefficient of z^n in

$$\exp\left(\gamma \sum_{k=1}^{\infty} \frac{m_k}{k} z^k\right).$$

Clearly, the latter and the former are two ways to compute the same number.

Step 2. Next, let us assume that $m_1, m_2, m_3, \ldots$ are moments of a compactly supported probability measure μ, i.e., there exists $r > 0$ and a probability measure μ supported inside $[-r, r]$, such that

$$m_k = \int_{-r}^{r} x^k \mu(dx), \qquad k = 1, 2, \ldots. \tag{7.7}$$

As it is true for any probability measure, μ can be approximated by discrete measures with atoms of weight $1/N$ as $N \to \infty$, and we can choose these measures to be supported inside $[-r, r]$. Let us fix such an approximation for μ, that is, we choose real numbers $-r \le a_1^{(N)} \le a_2^{(N)} \le \cdots \le a_N^{(N)} \le r$, such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \delta_{a_i^{(N)}} = \mu. \qquad \text{(Weak convergence of measures.)}$$

In particular, this implies

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N}(a_i^{(N)})^k = m_k, \qquad k = 1, 2, \dots.$$

Let μ_N be the Dirac delta-mass at the point $(a_1^{(N)}, \dots, a_N^{(N)})$—this is a probability measure on $\mathbb{R}^N$. Then the measures μ_N satisfy the Law of Large Numbers in the sense of Definition 3.1 with sequence m_k given by (7.7). The BGF of the measure μ_N is $B_{(a_1^{(N)},\dots,a_N^{(N)})}(x_1, \dots, x_N;\, \theta)$. Hence, Theorem 3.8 yields that

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \frac{\partial^l}{\partial y^l} \ln B_{(a_1^{(N)},\dots,a_N^{(N)})}(y, 0^{N-1};\, \theta)\Big|_{y=0} = (l-1)!\cdot\kappa_l, \tag{7.8}$$

where $\{\kappa_l\}_{l\ge 1} = T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\ge 1})$.

On the other hand, note that the formulas (7.5) and (7.6) have a limit in the regime $N\to\infty, \theta\to 0, \theta N\to\gamma$, which reads:

$$\lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} B_{(a_1^{(N)},\dots,a_N^{(N)})}(y, 0^{N-1}; \theta) = \sum_{k=0}^{\infty}\frac{c_k}{(\gamma)_k}y^k, \tag{7.9}$$

where the numbers c_k are found from the following Taylor series expansion:

$$\begin{aligned}\sum_{k=0}^{\infty} c_k z^k &= \lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \prod_{i=1}^{N}(1-a_i^{(N)}z)^{-\theta} = \lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \exp\left[-\theta\sum_{i=1}^{N}\ln\left(1-a_i^{(N)}z\right)\right]\\ &= \lim_{\substack{N\to\infty,\,\theta\to 0\\ \theta N\to\gamma}} \exp\left[\theta N\sum_{k=1}^{\infty}\frac{z^k}{k}\frac{1}{N}\sum_{i=1}^{N}(a_i^{(N)})^k\right] = \exp\left[\gamma\sum_{k=1}^{\infty}\frac{m_k z^k}{k}\right].\end{aligned} \tag{7.10}$$

Because $|a_i^{(N)}| \le r$ for all $1\le i\le N$, for any $\varepsilon > 0$ the convergence in (7.10) is uniform over $|z|\le r^{-1}-\varepsilon$. We claim that the convergence in (7.9) is uniform over y in compact subsets of $\mathbb{C}$. Indeed, the term-by-term convergence of (7.5) to (7.9) is evident from (7.10), while a tail bound on the series can be obtained from a uniform N-independent bound on the coefficients $c_k = c_k^{(N)}$ in (7.5), (7.6) of the form $|c_k^{(N)}| \le C\cdot r^{-k}$ for $C > 0$, which follows from the Cauchy integral formula applied to (7.6).

Comparing (7.8) with (7.9) and noting that uniform convergence of analytic functions implies convergence of their derivatives, we conclude that

$$\exp\left(\sum_{l=1}^{\infty}\frac{\kappa_l}{l}y^l\right) = \sum_{k=0}^{\infty}\frac{c_k}{(\gamma)_k}y^k. \tag{7.11}$$

The last identity together with (7.10) give (3.8).

Step 3. It remains to study the case when $\{m_k\}_{k\ge 1}$ is an arbitrary sequence of real numbers, rather than a sequence of moments of a compactly supported probability measure μ as in (7.7). Note that the relation $\{\kappa_l\}_{l\ge 1} = T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\ge 1})$ is equivalent to saying that for certain polynomials Q^l, we have

$$\kappa_l = Q^l(m_1, m_2, \dots, m_l), \quad l = 1, 2, \dots.$$

On the other hand, (3.7) of Theorem 3.11 is equivalent to saying that for certain polynomials $\tilde{Q}^l$, we have

$$\kappa_l = \tilde{Q}^l(m_1, m_2, \dots, m_l), \quad l = 1, 2, \dots.$$

Hence, in order to prove Theorem 3.11, we need to show that the polynomials Q^l and $\tilde{Q}^l$ coincide for each $l = 1, 2, \dots$. Note that two polynomials in l variables coincide if and only if they coincide as functions on a non-empty open set $D \subset \mathbb{R}^l$, and this D can be chosen in an arbitrary way. Fix l and define:

$$D := \left\{ \left(\tfrac{1}{l}\sum_{i=1}^{l} d_i,\ \tfrac{1}{l}\sum_{i=1}^{l}(d_i)^2, \dots, \tfrac{1}{l}\sum_{i=1}^{l}(d_i)^l \right) \,\middle|\, d_1 < d_2 < \dots < d_l \right\} \subset \mathbb{R}^l.$$

The set D is an image of an open set $\{(d_1, \dots, d_l) \subset \mathbb{R}^l \mid d_1 < \dots < d_l\}$ under a smooth map with non-vanishing Jacobian (equal to $l^{-l}\prod_{i<j}(d_i - d_j)$). Hence, D is open. By Step 2 the polynomials Q^l and $\tilde{Q}^l$ coincide as functions on D, because each element in D is a moment sequence of the discrete probability measure with atoms $\frac{1}{l}$ at points $d_1, \dots, d_l$. Therefore, polynomials Q^l and $\tilde{Q}^l$ coincide. □

8. Limits of the Maps $T^{\gamma}_{\kappa\to m}$ and $T^{\gamma}_{m\to\kappa}$ as $\gamma \to 0$ and $\gamma \to \infty$

In this section we investigate the behavior of the γ-cumulants and γ-convolution as $\gamma \to 0$ and $\gamma \to \infty$. We will see that in the former case the conventional cumulants and conventional convolution appear, while in the latter case we link to the free probability counterparts.

8.1. $\gamma \to 0$ limit. Let us recall the definition of the classical cumulants.

Definition 8.1. Given a sequence of moments $\{m_k\}_{k\geq 1}$ we define the corresponding cumulants $\{c_l\}_{l\geq 1} = \tilde{T}_{m\to c}(\{m_k\}_{k\geq 1})$ through the identity for the generating functions:

$$C(z) = \ln(M(z)), \qquad \text{where} \quad M(z) := 1 + \sum_{k=1}^{\infty} \frac{m_k}{k!} z^k, \quad C(z) := \sum_{l=1}^{\infty} \frac{c_l}{l!} z^l. \tag{8.1}$$

In the opposite direction, given a sequence of cumulants $\{c_l\}_{l\geq 1}$, we define the corresponding sequence of moments $\{m_k\}_{k\geq 1} = \tilde{T}_{c\to m}(\{c_l\}_{l\geq 1})$ through a combinatorial formula:

$$m_k := \sum_{\pi = B_1 \sqcup \dots \sqcup B_m \in \mathscr{P}(k)} \prod_{i=1}^{m} c_{|B_i|}, \quad k = 1, 2, \dots. \tag{8.2}$$

The two definitions (8.1) and (8.2) are well-known to be equivalent and the maps $\tilde{T}_{m\to c}$ and $\tilde{T}_{c\to m}$ are inverse to each other, see, e.g., [MS17, Sections 1.1–1.2].

Theorem 8.2. *Given a sequence of numbers $\{m_k\}_{k\geq 1}$, let*

$$\{\kappa^0_l\}_{l\geq 1} = \lim_{\gamma\to 0} T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\geq 1}), \qquad \{c_l\}_{l\geq 1} = \tilde{T}_{m\to c}(\{m_k\}_{k\geq 1}).$$

Then for each $l = 1, 2, \dots$, we have $\kappa^0_l = \frac{1}{(l-1)!} c_l$.

Remark 8.3. The second relation implies $\tilde{T}_{c\to m}(\{c_l\}_{l\geq 1}) = \{m_k\}_{k\geq 1}$. The first one means that if we let $T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\geq 1}) =: \{\kappa^{\gamma}_l\}_{l\geq 1}$ (the κ^{γ}_l's depend on γ), then $\lim_{\gamma\to 0}\kappa^{\gamma}_l = \kappa^0_l = \frac{1}{(l-1)!}c_l, l\geq 1$. As $T^{\gamma}_{\kappa\to m}$ is the inverse of $T^{\gamma}_{m\to\kappa}$, then $\{m_k\}_{k\geq 1} = T^{\gamma}_{\kappa\to m}(\{\kappa^{\gamma}_l\}_{l\geq 1})$. Observe that $T^{\gamma}_{\kappa\to m}(\{\kappa^{\gamma}_l\}_{l\geq 1})$ is a sequence where each entry is a polynomial (the one from Theorem 3.10) in the variables $\kappa^{\gamma}_l, l\geq 1$, and the coefficients of these polynomials have limits as $\gamma\to 0$, thus $\{m_k\}_{k\geq 1} = \lim_{\gamma\to 0} T^{\gamma}_{\kappa\to m}(\{\kappa^{\gamma}_l\}_{l\geq 1}) = \lim_{\gamma\to 0} T^{\gamma}_{\kappa\to m}(\{\frac{1}{(l-1)!}c_l\}_{l\geq 1})$. Hence we have two equations for $\{m_k\}_{k\geq 1}$:

$$\tilde{T}_{c\to m}(\{c_l\}_{l\geq 1}) = \{m_k\}_{k\geq 1}, \qquad \lim_{\gamma\to 0} T^{\gamma}_{\kappa\to m}(\{\tfrac{1}{(l-1)!}c_l\}_{l\geq 1}) = \{m_k\}_{k\geq 1}.$$

In the same fashion, one can show that these two relations imply the ones in the theorem. Hence the theorem is equivalent to the statement that these last two definitions of $\{m_k\}_{k\geq 1}$ coincide.

Proof of Theorem 8.2. We use the description of the map $T^{\gamma}_{m\to\kappa}$ of Theorem 3.10 and send $\gamma\to 0$ in the weight $W(\pi)$ of (3.4). Note that

$$\lim_{\gamma\to 0}\big[p(i)!(\gamma+p(i)+1)_{q(i)}\big] = (p(i)+q(i))!$$

According to definitions of Sect. 3.2, $p(i)+q(i)+1 = |B_i|$, i.e., the size of the ith block in π. Hence, the $\gamma\to 0$ limit of Theorem 3.10 gives

$$m_k = \sum_{\pi=B_1\sqcup\cdots\sqcup B_m\in\mathscr{P}(k)}\prod_{i=1}^{m}\Big[(|B_i|-1)!\,\kappa^0_{|B_i|}\Big], \qquad k=1,2,\dots. \tag{8.3}$$

Comparing with (8.2) and noting that the relations (8.3) uniquely determine $\{\kappa^0_l\}_{l\geq 1}$, we conclude that $\kappa^0_l = \frac{1}{(l-1)!}c_l$. □

As a corollary, we obtain the $\gamma\to 0$ behavior of the γ-convolution of Theorem 1.5.

Corollary 8.4. *Take two sequences of real numbers* $\{m^{\mathbf{a}}_k\}_{k\geq 1}$ *and* $\{m^{\mathbf{b}}_k\}_{k\geq 1}$. *Define*

$$\{\tilde{m}_k\}_{k\geq 1} := \lim_{\gamma\to 0}\Big[\{m^{\mathbf{a}}_k\}_{k\geq 1}\boxplus_{\gamma}\{m^{\mathbf{b}}_k\}_{k\geq 1}\Big].$$

Then with the agreement $m^{\mathbf{a}}_0 = m^{\mathbf{b}}_0 = 1$, *we have*

$$\tilde{m}_k = \sum_{s=0}^{k}\binom{k}{s}m^{\mathbf{a}}_s m^{\mathbf{b}}_{k-s}, \quad k=1,2,\dots. \tag{8.4}$$

Remark 8.5. Suppose that we are given two independent random variables $\mathbf{a}$ and $\mathbf{b}$, such that

$$m^{\mathbf{a}}_k = \mathbb{E}\mathbf{a}^k, \quad m^{\mathbf{b}}_k = \mathbb{E}\mathbf{b}^k, \quad k=1,2,\dots.$$

Then the formula (8.4) says that $\tilde{m}_k = \mathbb{E}(\mathbf{a}+\mathbf{b})^k$.

Proof of Corollary 8.4. By (1.11) we have for each $\gamma>0$

$$T^{\gamma}_{m\to\kappa}(\{m^{\mathbf{a}}_k\}_{k\geq 1}\boxplus_{\gamma}\{m^{\mathbf{b}}_k\}_{k\geq 1}) = T^{\gamma}_{m\to\kappa}(\{m^{\mathbf{a}}_k\}_{k\geq 1}) + T^{\gamma}_{m\to\kappa}(\{m^{\mathbf{b}}_k\}_{k\geq 1}).$$

Taking the limit $\gamma\to 0$ and using Theorem 8.2, we get

$$\tilde{T}_{m\to c}\big(\{\tilde{m}_k\}_{k\geq 1}\big) = \tilde{T}_{m\to c}\big(\{m^{\mathbf{a}}_k\}_{k\geq 1}\big) + \tilde{T}_{m\to c}\big(\{m^{\mathbf{b}}_k\}_{k\geq 1}\big).$$

Taking into account (8.1), the last identity is equivalent to (8.4). □

8.2. $\gamma \to \infty$ limit. Let us recall the definition of the free cumulants.

A set partition $\pi \in \mathscr{P}(k)$ is said to be *crossing* if there exist two distinct blocks $B, B' \in \pi$ and integers $1 \le x < y < z < w \le k$ such that $x, z \in B$ and $y, w \in B'$. A set partition $\pi \in \mathscr{P}(k)$ is said to be a *non-crossing set partition* if it is not crossing. In the notations of Sect. 3.2, the non-crossing set partitions are those which have $p(i) = 0$ for all i; in other words, there should be no crossings of roofs and legs. We denote the collection of all non-crossing set partitions of $[k]$ by $NC(k)$. For instance, out the seven partitions of [4] with two blocks, mentioned in Sect. 3.2, the following six are the non-crossing ones:

$$\{1\} \sqcup \{2,3,4\}, \quad \{1,3,4\} \sqcup \{2\}, \quad \{1,2,4\} \sqcup \{3\}, \quad \{1,2,3\} \sqcup \{4\},$$
$$\{1,2\} \sqcup \{3,4\}, \quad \{1,4\} \sqcup \{2,3\}.$$

Definition 8.6. Given a sequence of moments $\{m_k\}_{k\ge1}$ we define the corresponding free cumulants $\{r_l\}_{l\ge1} = \tilde{T}^{\infty}_{m\to r}(\{m_k\}_{k\ge1})$ through the identity for the generating functions:

$$G\big(R(z)+z^{-1}\big) = z, \qquad \text{where} \quad G(z) := z^{-1} + \sum_{k=1}^{\infty} \frac{m_k}{z^{k+1}}, \quad R(z) := \sum_{l=1}^{\infty} r_l z^{l-1}. \tag{8.5}$$

In the opposite direction, given a sequence of free cumulants $\{r_l\}_{l\ge1}$, we define the corresponding sequence of moments $\{m_k\}_{k\ge1} = \tilde{T}^{\infty}_{r\to m}(\{r_l\}_{l\ge1})$ through a combinatorial formula:

$$m_k := \sum_{\pi = B_1 \sqcup \cdots \sqcup B_m \in NC(k)} \prod_{i=1}^{m} r_{|B_i|}, \quad k = 1, 2, \ldots. \tag{8.6}$$

The relation $G\big(R(z)+z^{-1}\big) = z$ can be rewritten as $R(z) = G^{(-1)}(z) - z^{-1}$ and $R(z)$ defined in this way is called the *Voiculescu R-transform*. The equivalence between (8.5) and (8.6) is explained in [MS17, Sect. 2.4].

Theorem 8.7. *Given a sequence of numbers $\{m_k\}_{k\ge1}$ define*

$$\{\kappa_l^{\gamma}\}_{l\ge1} = T^{\gamma}_{m\to\kappa}(\{m_k\}_{k\ge1}), \qquad r_l = \lim_{\gamma\to\infty} \gamma^{l-1}\kappa_l^{\gamma}.$$

Then we have $\{r_l\}_{l\ge1} = \tilde{T}^{\infty}_{m\to r}(\{m_k\}_{k\ge1})$.

Remark 8.8. Just like in Remark 8.3, the theorem can be equivalently stated by saying that the following two definitions

$$\{m_k\}_{k\ge1} = \lim_{\gamma\to\infty} T^{\gamma}_{\kappa\to m}(\{\gamma^{1-l} r_l\}_{l\ge1}), \qquad \{m_k\}_{k\ge1} = \tilde{T}^{\infty}_{r\to m}(\{r_l\}_{l\ge1}),$$

of the sequence $\{m_k\}_{k\ge1}$ coincide.

Proof of Theorem 8.7. We use the reformulation of Remark 8.8 together with the description of the map $T^{\gamma}_{\kappa\to m}$ of Theorem 3.10 and send $\gamma \to \infty$ in the weight $W(\pi)$ of (3.4). Note that

$$\lim_{\gamma\to\infty} \big[\gamma^{-p(i)-q(i)} p(i)!(\gamma + p(i) + 1)_{q(i)}\big] = \begin{cases} 1, & \text{if } p(i) = 0, \\ 0, & \text{otherwise.} \end{cases}$$

According to the definitions of Sect. 3.2, $p(i)+q(i)+1=|B_i|$, so $\gamma^{1-|B_i|}r_{|B_i|}=\gamma^{-p(i)-q(i)}r_{|B_i|}$. Hence, the $\gamma\to\infty$ limit of Theorem 3.10 gives

$$m_k=\lim_{\gamma\to\infty}\sum_{\pi=B_1\sqcup\cdots\sqcup B_m\in\mathscr{P}(k)}\prod_{i=1}^{m}\left[\gamma^{-p(i)-q(i)}p(i)!(\gamma+p(i)+1)_{q(i)}r_{|B_i|}\right]$$
$$=\sum_{\pi=B_1\sqcup\cdots\sqcup B_m\in NC(k)}\prod_{i=1}^{m}r_{|B_i|},\quad k=1,2,\dots. \tag{8.7}$$

This is the same expression as (8.6). □

As a corollary, we obtain the $\gamma\to\infty$ behavior of the γ-convolution of Theorem 1.5.

Corollary 8.9. *Take two sequences of real numbers* $\{m_k^{\mathbf{a}}\}_{k\geq1}$ *and* $\{m_k^{\mathbf{b}}\}_{k\geq1}$. *Define*

$$\{\tilde{m}_k\}_{k\geq1}:=\lim_{\gamma\to\infty}\left[\{m_k^{\mathbf{a}}\}_{k\geq1}\boxplus_\gamma\{m_k^{\mathbf{b}}\}_{k\geq1}\right].$$

Then the sequence $\{\tilde{m}_k\}_{k\geq1}$ *is uniquely fixed by the identity*

$$\tilde{T}^{\infty}_{m\to r}\left(\{\tilde{m}_k\}_{k\geq1}\right)=\tilde{T}^{\infty}_{m\to r}\left(\{m_k^{\mathbf{a}}\}_{k\geq1}\right)+\tilde{T}^{\infty}_{m\to r}\left(\{m_k^{\mathbf{b}}\}_{k\geq1}\right) \tag{8.8}$$

Remark 8.10. Suppose that we are given two random variables $\mathbf{a}$ and $\mathbf{b}$, such that

$$m_k^{\mathbf{a}}=\mathbb{E}\mathbf{a}^k,\quad m_k^{\mathbf{b}}=\mathbb{E}\mathbf{b}^k,\quad k=1,2,\dots.$$

Then the formula (8.8) says that $\tilde{m}_k$ are the moments of the *free convolution* of $\mathbf{a}$ and $\mathbf{b}$.

Proof of Corollary 8.9. By (1.11) we have for each $\gamma>0$

$$T^{\gamma}_{m\to\kappa}\left(\{m_k^{\mathbf{a}}\}_{k\geq1}\boxplus_\gamma\{m_k^{\mathbf{b}}\}_{k\geq1}\right)=T^{\gamma}_{m\to\kappa}\left(\{m_k^{\mathbf{a}}\}_{k\geq1}\right)+T^{\gamma}_{m\to\kappa}\left(\{m_k^{\mathbf{b}}\}_{k\geq1}\right).$$

Taking the limit $\gamma\to\infty$ and using Theorem 8.7, we get (8.8). □

Acknowledgements. The authors would like to thank Alexey Bufetov and Greta Panova for helpful discussions. We are thankful to Maciej Dołęga for pointing us to the articles [D09,BDEG21], and for sending us a draft of a new version of the latter paper. We thank Octavio Arizmendi and Daniel Perales for directing us to their work [AP18]. We are grateful to two anonymous referees for their feedback. The work of V.G. was partially supported by NSF Grants DMS-1664619, DMS-1949820, by BSF grant 2018248, and by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin–Madison with funding from the Wisconsin Alumni Research Foundation.

9. Appendix: Law of Large Numbers for Fixed Temperature

The aim of this "Appendix" is to probe the possibility of a version of Theorem 3.8 in which $\theta>0$ is fixed and is not changing with N. The following claim is an analogue of *one direction* of Theorem 3.8.

Claim 9.1 (LLN for finite temperature). *Let $\{\mu_N\}_{N\geq 1}$ be a sequence of exponentially decaying probability measures on tuples $a_1 \leq \cdots \leq a_N$. For each N, let $G_{N;\theta}(x_1,\ldots,x_N) := G_\theta(x_1,\ldots,x_N;\mu_N)$ be the BGF of μ_N. Assume that the sequence $\{G_{N;\theta}\}_N$ satisfies the following conditions:*

(a) $\lim\limits_{N\to\infty} \dfrac{1}{N}\cdot \dfrac{\partial^l}{\partial x_i^l} \ln\left(G_{N;\theta}\right)\Big|_{x_1=\cdots=x_N=0} = (l-1)!\cdot c_l$, *for all* $l\in\mathbb{Z}_{\geq 1}$.

(b) $\lim\limits_{N\to\infty} \dfrac{1}{N}\cdot \dfrac{\partial}{\partial x_{i_1}}\cdots\dfrac{\partial}{\partial x_{i_r}} \ln\left(G_{N;\theta}\right)\Big|_{x_1=\cdots=x_N=0} = 0$, *for all* $i_1,\ldots,i_r\in\mathbb{Z}_{\geq 1}$ *such that the set* $\{i_1,\ldots,i_r\}$ *contains at least two distinct indices.*

Then the sequence $\{\mu_N\}_{N\geq 1}$ satisfies the following LLN (compare to Definition 3.1): there exist real numbers $\{m_k\}_{k\geq 1}$ such that for any $s=1,2,\ldots$ and any $k_1,\ldots,k_s\in\mathbb{Z}_{\geq 1}$, we have

$$\lim_{N\to\infty} \mathbb{E}_{\mu_N} \prod_{i=1}^{s}\left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{a_i}{N}\right)^{k_i}\right) = \prod_{i=1}^{s} m_{k_i}.$$

If this occurs, $\{c_l\}_{l\geq 1}$ and $\{m_k\}_{k\geq 1}$ are related by either

$$m_k = \sum_{\pi\in NC(k)}\ \prod_{B\in\pi}\left(\theta^{|B|-1}c_{|B|}\right),\quad k=1,2,\ldots,$$

or, equivalently,

$$m_k = \frac{1}{k+1}\cdot[z^{-1}]\left(\left(z^{-1}+\sum_{l=1}^{\infty}\theta^{l-1}c_l z^{l-1}\right)^{k+1}\right),\quad k\in\mathbb{Z}_{\geq 1}.$$

We do not present a proof of the claim, but probably one can prove it with the same techniques that we have used in Sect. 5. At $\theta=1$, another approach is by a degeneration of [BuG15, Theorem 5.1], see also [MN18]. This claim would prove the one-sided implication

conditions (a) and (b) (fixed θ version of LLN-appropriateness) $\Longrightarrow$ LLN-satisfaction.

Based on our Theorem 3.8, on [BuG19, Theorem 2.6] (which studies the CLT at fixed $\theta=1$), and on the classical theorem that relates the weak convergence of measures to the convergence of their characteristic functions, the reader may be inclined to believe that the reverse implication is also true and that this kind of "if and only if" results are always expected.

However, this turns out to be wrong. The naive analogue of Theorem 3.8 does not hold for fixed θ: the "expected if and only if statement" is false. Here is a counter-example.

Let us consider a sequence of probability measures μ_N such that a random μ_N-distributed vector $a_1\geq\cdots\geq a_N$ has $a_1=\cdots=a_N$ almost surely and this common value a is distributed according to a Gaussian measure of mean 0 and variance N. In this case, the random variable

$$p_k^N = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{a_i}{N}\right)^k = \left(\frac{a}{N}\right)^k$$

is distributed as the k-th power of a Gaussian random variable of mean 0 and variance $1/N$. Consequently, the sequence $\{\mu_N\}_N$ satisfies a LLN, and all m_k's are equal to zero. On the other hand, by using $B_{(a,\dots,a)}(x_1,\dots,x_N;\theta) = \exp(a\sum_{i=1}^N x_i)$, it follows that the BGF of μ_N equals

$$G_{N;\theta}(x_1,\dots,x_N) = \int_{-\infty}^{\infty} \frac{e^{-a^2/(2N)}}{\sqrt{2\pi N}} B_{(a,\dots,a)}(x_1,\dots,x_N;\theta)da = \exp\left(\frac{N}{2}\left(\sum_{i=1}^N x_i\right)^2\right).$$

The sequence $\{G_{N;\theta}\}_N$ of BGFs then satisfies

$$\lim_{N\to\infty} \frac{1}{N}\cdot \frac{\partial^2}{\partial x_1 \partial x_2} \ln(G_{N;\theta})\Big|_{x_1=\dots=x_N=0} = 1,$$

therefore contradicting condition (b) from Claim 9.1.

A more refined question is whether *some* "if and only if for LLN" statement holds, if one modifies somehow the conditions (a) and (b) of Claim 9.1. Based on small calculations (obtained when trying to reverse-engineer the proof of Theorem 3.8), it is plausible that the answer is yes.

Indeed, based on Proposition 2.11, we must study the limits of the expressions

$$N^{-|\lambda|-\ell(\lambda)} \cdot \left[\prod_{i=1}^{\ell(\lambda)} \mathcal{P}_{\lambda_i}\right] G_{N;\theta}, \tag{9.1}$$

where λ ranges over the set of all partitions of a given size k. We have performed calculations for $k = 2, 3$; they indicate that the conditions on second-order derivatives in (a) and (b) from Claim 9.1 should be replaced by:

- $\lim_{N\to\infty} \frac{1}{N}\left\{\frac{\partial^2}{\partial x_1^2} - \frac{\partial^2}{\partial x_1\partial x_2}\right\} \ln(G_{N;\theta})\Big|_{x_1=\dots=x_N=0} = \theta^{-1}\cdot c_2,$
- $\lim_{N\to\infty} \frac{1}{N^2} \frac{\partial^2 \ln(G_{N;\theta})}{\partial x_1 \partial x_2}\Big|_{x_1=\dots=x_N=0} = 0,$

and the conditions on third-order derivatives should be replaced by:

- $\lim_{N\to\infty} \frac{1}{N}\left\{\frac{1}{2}\cdot\frac{\partial^3}{\partial x_1^3} - \frac{3}{2}\cdot\frac{\partial^3}{\partial x_1^2\partial x_2} + \frac{\partial^3}{\partial x_1\partial x_2\partial x_3}\right\} \ln(G_{N;\theta})\Big|_{x_1=\dots=x_N=0} = \theta^{-2}\cdot c_3,$
- $\lim_{N\to\infty} \frac{1}{N^2}\left\{\frac{\partial^3}{\partial x_1^2\partial x_2} - \frac{\partial^3}{\partial x_1\partial x_2\partial x_3}\right\} \ln(G_{N;\theta})\Big|_{x_1=\dots=x_N=0} = 0,$
- $\lim_{N\to\infty} \frac{1}{N^3} \frac{\partial^3 \ln(G_{N;\theta})}{\partial x_1\partial x_2\partial x_3}\Big|_{x_1=\dots=x_N=0} = 0.$

These relations are much more involved than conditions (a) and (b) from Claim 9.1, or than the conditions from Definition 3.3. What should be the correct "if and only if" relations for $k > 3$? This is an interesting open question for future research.

References

[A20] Ahn, A.: Airy Point Process via Supersymmetric Lifts. arXiv:2009.06839

[AB19] Akemann, G., Byun, S.-S.: The high temperature crossover for general 2D Coulomb gases. J. Stat. Phys. **175**(6), 1043–1065 (2019)

[ABG12] Allez, R., Bouchaud, J.-P., Guionnet, A.: Invariant Beta ensembles and the Gauss–Wigner crossover. Phys. Rev. Lett. **109**(9), 094–102 (2012)

[ABMV13] Allez, R., Bouchaud, J.-P., Majumdar, S.N., Vivo, P.: Invariant β-Wishart ensembles, crossover densities and asymptotic corrections to the Marcenko–Pastur law. J. Phys. A Math. Theor. **46**(1), 015001 (2013)

[AD14] Allez, R., Dumaz, L.: From Sine kernel to Poisson statistics. Electron. J. Probab. **19**(114), 1–25 (2014). arXiv:1407.5402

[A91] Anderson, G.W.: A Short Proof of Selberg's Generalized Beta Formula. In Forum Mathematicum **3**(3), 415418 (1991)

[AHV20] Andraus, S., Hermann, K., Voit, M.: Limit theorems and soft edge of freezing random matrix models via dual orthogonal polynomials, to appear in Journal of Mathematical Physics. arXiv:2009.01418

[A17] Anker, J.-P.: An introduction to Dunkl theory and its analytic aspects. Analytic, Algebraic and Geometric Aspects of Differential Equations. Birkhäuser, Cham, pp. 3–58 (2017)

[AP18] Arizmendi, O., Perales, D.: Cumulants for finite free convolution. J. Comb. Theory Ser. A **155**, 244–266 (2018)

[AW84] Askey, R., Wimp, J.: Associated Laguerre and Hermite polynomials. Proc. R. Soc. Edinb. Sect. A Math. **96**(1–2), 15–37 (1984)

[AN19] Assiotis, T., Najnudel, J.: The boundary of the orbital beta process. arXiv:1905.08684

[BAG97] Ben Arous, G., Guionnet, A.: Large deviations for Wigner's law and Voiculescu's non-commutative entropy. Probab. Theory Relat. Fields **108**, 517–542 (1997)

[BP15] Benaych-Georges, F., Péché, S.: Poisson statistics for matrix ensembles at large temperature. J. Stat. Phys. **161**(3), 633–656 (2015)

[BV95] Bercovici, H., Voiculescu, D.: Superconvergence to the central limit and failure of the Cramér theorem for free random variables. Probab. Theory Relat. Fields **102**, 215–222 (1995)

[BoG15] Borodin, A., Gorin, V.: General β-Jacobi Corners Process and the Gaussian Free Field. Commun. Pure Appl. Math. **68**(10), 1774–1844 (2015)

[BDEG21] Bożejko, M., Dołęga, M., Ejsmont, W., Gal, Ś.R.: Reflection length with two parameters in the asymptotic representation theory of type B/C and applications. arXiv:2104.14530

[BuG15] Bufetov, A., Gorin, V.: Representations of classical Lie groups and quantized free convolution. Geom. Funct. Anal. **25**, 763–814 (2015)

[BuG18] Bufetov, A., Gorin, V.: Fluctuations of particle systems determined by Schur generating functions. Adv. Math. **338**, 702–781 (2018)

[BuG19] Bufetov, A., Gorin, V.: Fourier transform on high-dimensional unitary groups with applications to random tilings. Duke Math. J. **168**(13), 2559–2649 (2019)

[Cu19] Cuenca, C.: Universal behavior of the corners of orbital beta processes. Int. Math. Res. Not. **19**, 14761–14813 (2021)

[Co03] Collins, B.: Moments and cumulants of Polynomial random variables on unitary groups, the Itzykson Zuber integral and free probability. Int. Math. Res. Not. **2003**(17), 953–982 (2003)

[CS06] Collins, B., Sniady, P.: Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. Commun. Math. Phys. **264**, 773–795 (2006)

[D09] Drake, D.: The combinatorics of associated Hermite polynomials. Eur. J. Comb. **30**(4), 1005–1021 (2009)

[DE06] Dumitriu, I., Edelman, A.: Global spectrum fluctuations for the β-Hermite and β-Laguerre ensembles via matrix models. J. Math. Phys. **47**, 063302 (2006)

[D89] Dunkl, C.: Differential-difference operators associated to reflection groups. Trans. Am. Math. Soc. **311**(1), 167–183 (1989)

[DS15] Duy, K.T., Shirai, T.: The mean spectral measures of random Jacobi matrices related to Gaussian beta ensembles. Electron. Commun. Probab. **20** (2015)

[EM10] Etingof, P., Ma, X.: Lecture notes on Cherednik algebras. Preprint. arXiv:1001.0432 (2010)

[FF16] Faraut, J., Fourati, F.: Markov–Krein transform. Colloq. Math. **144**(1), 137–156 (2016)

[F71] Feller, W.: An Introduction to Probability: Theory and its Applications, vol. II, 2nd edn. Wiley (1971)

[F21] Forrester, P.J.: High-low temperature dualities for the classical β-ensembles. arXiv:2103.11250

[FM21] Forrester, P.J., Mazzuca, G.: The classical β-ensembles with β proportional to $1/N$: from loop equations to Dyson's disordered chain. arXiv:2102.09201

[FD16] Fyodorov, Y., Le Doussal, P.: Moments of the position of the maximum for GUE characteristic polynomials and for log-correlated Gaussian processes. J. Stat. Phys. **164**, 190–240 (2016)
[GK20] Gorin, V., Kleptsyn, V.: Universal objects of the infinite beta random matrix theory. arXiv:2009.02006
[GM20] Gorin, V., Marcus, A.W.: Crystallization of random matrix orbits. Int. Math. Res. Not. **2020**(3), 883–913 (2020)
[GP15] Gorin, V., Panova, G.: Asymptotics of symmetric polynomials with applications to statistical mechanics and representation theory. Ann. Probab. **43**(6), 3052–3132 (2015). arXiv:1301.0634
[GSh15] Gorin, V., Shkolnikov, M.: Multilevel Dyson Brownian motions via Jack polynomials. Probab. Theory Relat. Fields **163**(3), 413–463 (2015). arXiv:1401.5595
[GSu18] Gorin, V., Sun, Y.: Gaussian fluctuations for products of random matrices, to appear in American Journal of Mathematics. arXiv:1812.06532 (2018)
[GuK02] Guhr, T., Kohler, H.: Recursive construction for a class of radial functions. I. Ordinary space. J. Math. Phys. **43**(5), 2707–2740 (2002)
[HL21] Hardy, A., Lambert, G.: CLT for Circular beta-Ensembles at high temperature. J. Funct. Anal. **280**(7), 108869 (2021)
[J98] Johansson, K.: On fluctuations of eigenvalues of random Hermitian matrices. Duke Math. J. **91**(1), 151–204 (1998)
[Hu18] Huang, J.: Law of large numbers and central limit theorems by jack generating functions. Adv. Math. **380**, 107545 (2021)
[K98] Kerov, S.: Interlacing measures. American Mathematical Society Translations, 35–84 (1998)
[K03] Kerov, S.: Asymptotic Representation Theory of the Symmetric Group and its Applications in Analysis. American Mathematical Society, Providence (2003)
[KS09] Killip, R., Stoiciu, M.: Eigenvalue statistics for CMV matrices: from Poisson to clock via random matrix ensembles. Duke Math. J. **146**, 361–399 (2009)
[K97] Kirillov, A., Jr.: Lectures on affine Hecke algebras and Macdonald's conjectures. Bull. Am. Math. Soc. **34**(3), 251–292 (1997)
[KT] Knutson, A., Tao, T.: Honeycombs and sums of Hermitian matrices. Not. Am. Math. Soc. **48**(2), 175–186 (2001)
[L70] Lukacs, E.: Characteristic Functions, 2nd edn. Griffin (1970)
[M98] Macdonald, I.G.: Symmetric Functions and Hall Polynomials. Oxford University Press (1998)
[Ma18] Marcus, A.W.: Polynomial convolutions and (finite) free probability, Preprint (2018)
[MN18] Matsumoto, S., Novak, J.: A moment method for invariant ensembles. Electron. Res. Announc., **25**, 60
[M19] Matveev, K.: Macdonald-positive specializations of the algebra of symmetric functions: Proof of the Kerov conjecture. Ann. Math. **189**(1), 277–316 (2019)
[MP21] Mergny, P., Potters, M.: Rank one HCIZ at high temperature: interpolating between classical and free convolutions. Preprint. arXiv:2101.01810 (2021)
[MS17] Mingo, J.A., Speicher, R.: Free Probability and Random Matrices, vol. 35. Springer, New York (2017)
[NT18] Nakano, F., Trinh, K.D.: Gaussian beta ensembles at high temperature: eigenvalue fluctuations and bulk statistics. J. Stat. Phys. **173**(2), 296–321 (2018)
[NTT21] Nakano, F., Trinh, H.D., Trinh, K.D.: Limit theorems for moment processes of beta Dyson's Brownian motions and beta Laguerre processes. arXiv:2103.09980
[Ne03] Neretin, Y.A.: Rayleigh triangles and non-matrix interpolation of matrix beta integrals. Sbornik: Math. **194**(4), 515–540 (2003)
[NS06] Nica, A., Speicher, R.: Lectures on the Combinatorics of Free Probability, vol. 13. Cambridge University Press (2006)
[OO97] Okounkov, A., Olshanski, G.: Shifted Jack polynomials, binomial formula, and applications. Math. Res. Lett. **4**, 69–78 (1997)
[OV96] Olshanski, G., Vershik, A.: Ergodic unitarily invariant measures on the space of infinite Hermitian matrices, In: Contemporary Mathematical Physics. F. A. Berezin's memorial volume. Amer. Math. Transl. Ser. 2, vol. 175 R. L. Dobrushin et al., (eds), pp. 137–175 (1996). arXiv:math/9601215
[O93] Opdam, E.M.: Dunkl operators, Bessel functions and the discriminant of a finite Coxeter group. Compos. Math. **85**(3), 333–373 (1993)
[Pa18] Pakzad, C.: Poisson statistics at the edge of Gaussian beta-ensembles at high temperature. arXiv:1804.08214 (2018)
[Pi91] Pickrell, D.: Mackey analysis of infinite classical motion groups. Pac. J. Math. **150**, 139–166 (1991)
[R03a] Rösler, M.: A positive radial product formula for the Dunkl kernel. Trans. Am. Math. Soc. **355**(6), 2413–2438 (2003). arXiv:math/0210137

[R03b] Rösler, M.: Dunkl Operators: Theory and Applications. In: Koelink, E., Van Assche, W. (eds) Orthogonal Polynomials and Special Functions. Lecture Notes in Mathematics, vol. **1817**, (2003). Springer, Berlin

[R91] Rudin, W.: Functional Analysis. International Series in Pure and Applied Mathematics, vol. 8, 2nd edn. McGraw-Hill Science/Engineering/Math, New York, NY (1991)

[STJ20] Shlyakhtenko, D., Tao, T.: (with an appendix by David Jekel), Fractional free convolution powers, arXiv:2009.01882

[SZ20] Sokal, A.D., Zeng, J.: Some multivariate master polynomials for permutations, set partitions, and perfect matchings, and their continued fractions. arXiv:2003.08192

[S19] Spohn, H.: Generalized gibbs ensembles of the classical toda chain. J. Stat. Phys. **180**, 4–22 (2020)

[S89] Stanley, R.P.: Some combinatorial properties of jack symmetric functions. Adv. Math. **77**, 76–115 (1989)

[T02] Trimèche, K.: Payley–Wiener Theorems for the Dunkl Tramsform and Dunkl Translation Operators. Integral Transform. Spec. Funct. **13**(1), 17–38 (2002)

[TT19] Trinh, H.D., Trinh, H.D.: Beta Laguerre ensembles in global regime. Osaka J. Math. **58**, 435–450 (2021)

[TT20] Trinh, H.D., Trinh, K.D.: Beta Jacobi ensembles and associated Jacobi polynomials. arXiv:2005.01100

[V91] Voiculescu, D.: Limit laws for random matrices and free products. Invent. Math. **104**, 201–220 (1991)

[W12] Weyl, H.: Das asymptotische Verteilungsgesetz der Eigenwerte lineare partieller Differentialgleichungen. Math. Ann. **71**, 441–479 (1912)